Philipps Universität Marburg

Gutzwiller variational wave function for a two-orbital Hubbard model on a square lattice

Dissertation
zur
Erlangung des Doktorgrades
der Naturwissenschaften
(Dr. rer. nat.)

dem

Fachbereich Physik
der Philipps-Universität Marburg

vorgelegt von

Kevin Torben Graf zu Münster
Freiherr von Grothaus

aus Borchel bei Rotenburg/Wümme

Marburg/Lahn 2015

Vom Fachbereich Physik der Philipps-Universität
als Dissertation angenommen am .

Erstgutachter: Prof. Dr. Florian Gebhard
Fachbereich Physik
Arbeitsgruppe Vielteilchenphysik

Zweitgutachter: Prof. Dr. Peter Lenz
Fachbereich Physik
Arbeitsgruppe Komplexe Systeme

Tag der mündlichen Prüfung:

Hochschulkennziffer 1180

Zusammenfassung

Die elektronischen Eigenschaften vieler Materialien können durch eine effektive Einteilchentherorie beschrieben werden. Die Fermiflüssigkeitstheorie bietet zum Beispiel eine phänomenologische Beschreibung, wie die Wechselwirkung zwischen den Elektronen die Eigenschaften eines Fermigases beeinflusst [1]. Ausgehend von einem Grundzustand freier Teilchen, werden die wechselwirkenden Teilchen durch freie Teilchen mit einer renormalisierten Bandstruktur beschrieben. So verändert die Wechselwirkung lediglich die Eigenschaften der freien Teilchen in der Nähe der Fermikante. Auch fortgeschrittenere Methoden, wie die Dichtefunktionaltheorie, bauen auf einer effektiven Einteilchentheorie auf. Mit ihrer Hilfe kann man numerisch die renormalisierten Bandstrukturen berechnen und sogar eine ab-initio Berechnung der Gitterstrukur durchführen. Ist die Wechselwirkung jedoch stark genug, um ausgeprägte Korrelationseffekte zwischen den Elektronen zu bewirken, so können diese nicht länger in einem Einteilchenbild beschrieben werden.

Hubbard-Modelle bieten eine Möglichkeit, das Verhalten von stark korrelierten Elektronen auf einem Kristallgitter zu untersuchen. In diesen Modellen, wird der elektronische Hamilton-Operator mit Hilfe der „tight binding approximation" vereinfacht, siehe z. B. [2]. Die Bewegung der Elektronen wird durch Tunnelamplituden beschrieben, die es den Elektronen ermöglichen, auf einen benachbarten Gitterplatz zu hüpfen. Zusätzlich wird eine abstoßende Wechselwirkung zwischen den Elektronen, die sich auf dem gleichen Gitterplatz befinden, eingeführt. Die Hubbard-Modelle bieten die Möglickeit, ein weites Spektrum an elektronischen Phasen wie (Anti)-Ferromagnetismus, Supraleitung oder den Mott-Isulator zu studieren. Im Allgemeinen ist es jedoch nicht möglich, eine exakte Lösung des Grundzustandes zu finden. Numerische Näherungsverfahren erweisen sich als sehr aufwändig und können meist nur für kleine, endliche Systeme verwendet werden.

In dieser Arbeit wird der Grundzustand des korrelierten Systems durch die Gutzwiller-Variationswellenfunktion beschrieben [3]. Die Gutzwiller-Variationswellenfunktion baut auf einem Grundzustand eines Systems unabhängiger Teilchen auf, die sich frei im Gitter bewegen können. Dieser Grundzustand besteht aus einer Linearkombination von Zuständen, in denen sich die Teilchen statistisch über alle Gitterplätze verteilen. Um jedoch energetisch ungünstige Zustände mit vielen Mehrfachbesetzungen zu vermeiden, wird mit Hilfe des Gutzwiller-Korrelators das Gewicht dieser Zustände reduziert. Dadurch werden lokale Korrelationen in die Wellenfunktion eingebaut. In einem weiteren Schritt wird der zu Grunde gelegte Einteilchengrundzustand optimiert. Die exakte Auswertung der Gutzwiller-Wellenfunktion ist wei-

terhin mit einem erheblichen numerischem Aufwand verbunden. Nur der eindimensionale Einband-Fall [4–7] und der unendlich-dimensionale Fall [8–10], welcher als Gutzwiller-Näherung bekannt ist, lassen sich analytisch berechnen.

Im Allgemeinen kann die Gutzwiller-Wellenfunktion durch eine diagrammatische Entwicklung ausgewertet werden. Diese Entwicklung ähnelt der allgemein bekannten diagrammatischen Entwicklung der Green-Funktionen in der Festkörperphysik. Im Spezialfall eines unendlich-dimensionalen Gitters ergibt sich, dass sämtliche nicht trivialen Diagramme aufgrund des Skalenverhaltens der Einteilchendichtematrix verschwinden. Für die Anwendung auf ein endlich-dimensionales System müssen sämtliche Diagramme in Betracht gezogen werden. Dabei zeigt sich, dass die zusätzlichen Terme starke Auswirkungen auf die Optimierung des Einteilchengrundzustandes haben können. So wurde in [13] gezeigt, dass die Gutzwiller-Wellenfunktion in zwei Dimensionen zu einer Phase mit gebrochener Gittersymmetrie führen kann.

Die Gutzwiller-Variationswellenfunktion wird für ein ein Mehrband-Hubbard-Modell hergeleitet und auf ein Zweiband-Modell angewendet. Das Zweiband-Modell beschreibt Elektronen auf einem quadratischem Gitter mit lokalen p_x-p_y (bzw. d_{xy}-d_{yz}) Atomorbitalen. Es werden alle Tunnelamplituden zu den nächsten sowie übernächsten Nachbarn berücksichtigt. Die Gittersymmetrien werden in den Tunnelamplituden und der lokale Wechselwirkung der Elektronen berücksichtigt. Anschließend werden der ferromagnetischen Phasenübergang sowie Deformationen der Fermifläche, welche durch die Wechselwirkung der Teilchen induziert werden, untersucht. Dabei werden alle Diagramme mit bis zu zwei internen Knoten berechnet.

Ein Vergleich mit der Hartree-Fock-Theorie zeigt, dass die Gutzwiller-Wellenfunktion erst bei weitaus größeren Wechselwirkungen magnetisch wird. Zudem lässt sich erkennen, dass sich der Bereich, in dem ein ungesättigter Magnetismus auftritt, auf einen viel größeren Parameterbereich erstreckt. Dies lässt zum Beispiel einen einfacheren Abgleich der Modellparameter des Hubbard-Modells an experimentelle Daten zu. Die Deformationen der Fermifläche treten in dem Bereich auf, in dem die potentielle Energie von derselben Größenordnung wie die kinetische Energie ist. Die stärksten Deformationen können in der Nähe halber Bandfüllung beobachten werden. Es zeigt sich, dass die Deformationen sogar zu einer Änderung der Topologie der Fermifläche führen können. So wird gezeigt, dass die Korrelationen starken Einfluss auf die Form der Fermifläche nehmen können. Die optimierte Fermifläche kann zum Beispiel als Ausgangspunkt für eine Fermiflüssigkeitstheorie verwendet werden.

Contents

List of Figures

Chapter 1

Introduction

1.1 Scope of the work

The electronic properties of many materials can be described by an effective theory of independent electrons. Fermi liquid theory provides an analytical description of the effects of the interaction between the electrons on the independent particle picture. For a Fermi gas ground state, the energy can be written as a sum over the energy of all occupied single-particle states. Fermi liquid theory starts from modified single-particle energies [1]. In this phenomenological picture, the electrons behave like independent-particles with a renormalized band structure which is affected by the particle interactions. More sophisticated theories like density functional theory also employ an independent-particle picture. They allow for a numerical evaluation of the renormalized band structures and even an ab-initio calculation of the lattice structure. However, these theories do not satisfactorily describe the properties of strongly correlated systems, in which the electrons' interaction energy is of the same order as the kinetic energy.

In order to study the principle behavior of such strongly-correlated lattice systems, Hubbard-type models can be used. In these models, the tight-binding approximation is used as a starting point. See, e.g., [2] for an introduction. In this approximation, the electrons are supposed to occupy atomic orbitals on the lattice positions. They can tunnel between nearest (and next-nearest) neighbors so that they form narrow bands. Their Coulomb interaction is neglected if the electrons do not occupy the same site. It is known that if the local interactions become very strong, the electrons of such model system can display various ordered phases such as (anti-)ferromagnetism, superconductivity or a Mott insulating phase. In general, the characteristics of the phases strongly depends on the model specifications and on the dimension and structure of the lattice. Unfortunately, it is notoriously difficult to treat even the simplest Hubbard-type models because of the complexity of the full many-body quantum system.

In this work, we employ the Gutzwiller wave function approach, see, e.g., [3], to treat the electronic correlations. The Gutzwiller variational ground state starts from the independent-particle picture where the electrons are distributed over all lattice

sites to optimize the kinetic energy. This statistical distribution leads to atomic configurations that are energetically unfavorable for the Hubbard interaction. In the Gutzwiller wave function, the weight of such configurations is reduced with the help of the Gutzwiller correlator.

In this way, we can include local correlations into the ground state of the noninteracting system. However, the evaluation of Gutzwiller-correlated wave functions again poses a very difficult many-particle problem. Only the Gutzwiller-correlated single-band Fermi sea can be calculated exactly in one dimension [4–7]. A multi-band Gutzwiller-correlated state can be evaluated in the limit of infinite dimensions, where the result of the so-called Gutzwiller approximation becomes exact [8–10]. Since the Gutzwiller method in high dimensions is applicable to multi-band models, it could be applied to real materials, e.g., to describe the itinerant ferromagnetism of nickel [11]. The authors started from an effective single-particle Hamiltonian which was obtained with a density functional calculation. Then, a Gutzwiller theory was used to improved the correct treatment of the electronic correlations. More recently, this 'Gutzwiller DFT' approach is combined in a self-consistent manner [12].

As in standard many-body theory, the evaluation of expectation values requires the calculation of diagrams to infinite order. The Gutzwiller correlator permits the setup of a diagrammatic formalism in such a way that, in the infinite dimensional limit, the scaling of the single-particle density matrix leads to a cancellation of all nontrivial diagrams. However, the evaluation of the action of the Gutzwiller correlator in infinite dimensions neglects important spatial correlations of the density matrix. It has been shown, for example, that the Gutzwiller wave function in two dimensions can lead to a broken spatial symmetry of the Fermi surface [13]. This effect cannot be described in the limit of infinite dimensions. In finite dimensions, the evaluation of the Gutzwiller wave function on a square lattice requires the evaluation of all diagrams. In this work, we derive the diagrammatic expansion of a two-orbital model on a square lattice. We compute the ground-state energy up to and including two internal vertices. As applications, we adress (i) the ferromagnetic phase transitions as a function of the band-filling, and (ii) the Fermi surface deformations induced by the interaction. We confirm preliminary findings that ferromagnetism is a phenomena of strongly correlated electrons. Moreover, we find that correlation-induced deformations of the Fermi surface can be substantial and that they can even change the Fermi surface topology.

1.2 Outline

This work is structured as follow. In the second chapter, we develop the diagrammatic analysis of multi-band Gutzwiller wave functions in finite dimensions. The derivation of this analysis does not depend on the model specifications or the dimensionality of the lattice. First, we give the necessary definitions and fix our notation. Next, we use Wicks's theorem to formulate our diagrammatic approach. Lastly, we derive the linked cluster theorem for the multi-band case so that a finite number of diagrams has to be calculated in each order of the expansion.

In the third chapter, we introduce a Hubbard model with p_x-p_y orbitals on a square lattice. We consider nearest and next-nearest neighbor hopping terms and focus on purely local interactions. Then, we examine the ferromagnetic phase transition and confirm that the Gutzwiller wave function predicts ferromagnetism for much larger interactions than the corresponding Hartree-Fock/Stoner theory. Then, we use the diagrammatic expansion to optimize the underling Fermi state. The variational approach shows that the electronic correlations can lead to strong deformation of the Fermi surface. In some cases, even the topology of the Fermi surface changes. This shows that a simple application of the Fermi liquid theory or any other theory which starts from the noninteracting Fermi surface will not be suitable to describe the system properly. A summary and an outlook conclude this work. Technical details are deferred to the appendix.

Chapter 2

Diagrammatic analysis

In the first section of this chapter, we introduce the multi-band Hubbard model and the Gutzwiller variational ground state, and fix the notation necessary for the discussion of multi-band problems. Then, we calculate expectation values with Gutzwiller wave functions in a perturbation series whose coefficients can be represented by diagrams. In standard Feynman-Dyson perturbation theory for interacting electrons, 'lines' and 'vertices' in the diagrammatic analysis are predefined. In the Gutzwiller approach, in contrast, we can arrange the expansion in such a way that all internal diagrams vanish in infinite dimensions. This also improves the convergence of the perturbation series in finite dimensions.

To illustrate the procedure, we address the expectation values for a local operator $\hat{O}_i$ on lattice site i and Gutzwiller wave function $|\Psi_G\rangle$,

$$\langle \hat{O}_i \rangle_G = \frac{\langle \Psi_G | \hat{O}_i | \Psi_G \rangle}{\langle \Psi_G | \Psi_G \rangle} \, .$$

Diagrams involve the external vertex on site i and n internal nodes in nth order of our perturbation expansion. The vertices are linked by lines representing the single-particle density matrix. Lowest order diagrams are shown in Fig. 2.1. A more detailed derivation will be given later in this chapter.

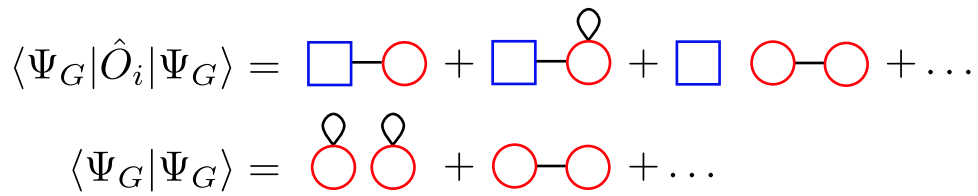


Figure 2.1: Diagrammatic representation of the numerator and denominator of the expectation value of an local operator $\hat{O}_i$ on the lattice site i. The blue square and the red circles gives the external and internal nodes, respectively. Black lines correspond to the single-particle density matrix.

In a first step, we eliminate all diagrams with local loops. To this end, we introduce the so-called Hartree-Fock operators which will be discussed in section 2.2. Fig. 2.2 illustrates the effect of working with Hartree-Fock operators in the series expansion.

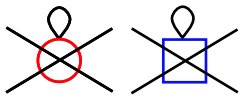


Figure 2.2: All nodes with internal lines cancel out after a redefinition of the nodes.

In section 2.3 we apply the linked cluster theorem so that all unconnected diagrams in the numerator cancel the denominator. Then, the nth order in pertubation theory can be expressed by a finite number of connected diagrams with n internal nodes. As a final step, in section 2.4, we get rid of all internal nodes that have only two lines, as shown in Fig. 2.3.

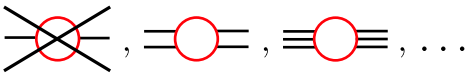


Figure 2.3: The internal nodes with only two lines will be eliminated by a gauge in the variational parameters.

The last step can be achieved by a gauge in the variational parameters. In infinite dimensions, the expectation value of the Hubbard Hamiltonian only involves external vertices, as shown in Fig. 2.4. In finite dimension, we must also consider the higher order diagrams with a finite number of internal nodes. Fig. 2.4 shows some second order diagrams for the interaction and the kinetic energy with nodes that connect to at least four lines.

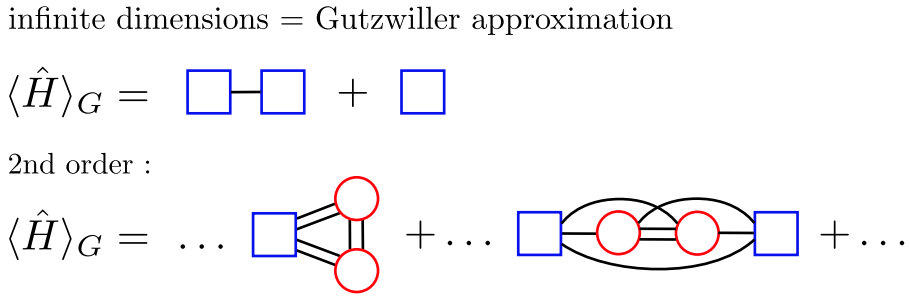


Figure 2.4: Diagrammatic analysis in infinite and finite dimensions.

2.1 Definitions and notation

To be more specific, we start with an introduction of the mathematical quantities and the notation used in this work. The notation closely follows Ref. [14].

2.1.1 Hamilton operator

We consider the Hubbard model with purely local interactions

$$\hat{H} = \hat{T} + \hat{U} = \sum_{i,j,\sigma,\sigma'} t_{ij}^{\sigma\sigma'} \hat{c}_{i,\sigma}^{\dagger} \hat{c}_{j,\sigma'} + \sum_{i,\sigma_1,\ldots,\sigma_4} U_{\sigma_1\sigma_2\sigma_3\sigma_4} \hat{c}_{i,\sigma_1}^{\dagger} \hat{c}_{i,\sigma_2}^{\dagger} \hat{c}_{i,\sigma_3} \hat{c}_{i,\sigma_4} \,. \tag{2.1.1}$$

Here, $\hat{c}^\dagger_{j,\sigma}$ and $\hat{c}_{j,\sigma'}$ are fermionic creation and annihilation operator, respectively. The site index is given by j and the spin-orbital index by σ. The hopping amplitudes are given by $t_{ij}^{\sigma\sigma'}$ and the interaction strength of the on-site interaction is given by the coefficients $U_{\sigma_1\sigma_2\sigma_3\sigma_4}$. The lattice indices run over all lattice sites of the lattice Λ. Periodic boundary conditions apply. Ultimately, we are interested in the thermodynamic limit in which the lattice size becomes infinite.

2.1.2 Local configuration space

The local configuration space of our fermionic system shall be restricted to n_B orbitals, so that the maximal number of possible local configurations is given by 2^{2n_B}. The fully occupied configuration is denoted by the full set Ω and the empty configuration by the empty set $\varnothing$. The local Hilbert space basis is given by

$$\tilde{\Omega} = \{\varnothing, \{\sigma_1\}, \{\sigma_2\}, \ldots, \{\sigma_n\}, \{\sigma_1, \sigma_n\}, \ldots, \Omega\backslash\sigma_n, \Omega\} \; . \tag{2.1.2}$$

The index I_m denotes an element of $\tilde{\Omega}$ and $|I_m\rangle$ denotes the state in which all orbitals of the configurations in I_m are occupied. We can define a product of local creation or annihilation operators by the introduction of the following symbols

$$\begin{aligned} \hat{C}^\dagger_I &= \prod_{\sigma\in I} \hat{c}^\dagger_\sigma = \hat{c}^\dagger_{\sigma_1} \ldots \hat{c}^\dagger_{\sigma_n} \quad i<j \rightarrow \sigma_i < \sigma_j \; , \\ \hat{C}_I &= \prod_{\sigma\in I} \hat{c}_\sigma = \hat{c}_{\sigma_1} \ldots \hat{c}_{\sigma_n} \quad i<j \rightarrow \sigma_i > \sigma_j \; . \end{aligned} \tag{2.1.3}$$

We define the standard ordering of the index σ_i as $u_1, d_1, u_2, d_2, \ldots$, where u and d define the spin index and the subindex gives the orbital index. The ordering of the multi-particle states is given by the lexicographical ordering of their subindexes. Here, we start with a comparison of the highest index and if they are equal move on to the next highest index, and proceed until we encounter the first unequal index.

When we pull a set of creation or annihilation operators out of a bigger set, we need to commute the operators until they are placed in the proper position. In this procedure the overall sign may change. We introduce the following symbols to indicate that we consider a sign change after a reordering procedure

$$\begin{aligned} &[\overrightarrow{I,J}]^+_\Sigma \quad \text{sign after splitting } \hat{C}^\dagger_{I\cup J} \rightarrow \hat{C}^\dagger_I \hat{C}^\dagger_J \; , \\ &[\overrightarrow{I,J}]^-_\Sigma \quad \text{sign after splitting } \hat{C}_{I\cup J} \rightarrow \hat{C}_I \hat{C}_J \; , \\ &[\overleftarrow{J,I}]^+_\Sigma \quad \text{sign after splitting } \hat{C}^\dagger_{I\cup J} \rightarrow \hat{C}^\dagger_J \hat{C}^\dagger_I \; , \\ &[\overleftarrow{J,I}]^-_\Sigma \quad \text{sign after splitting } \hat{C}_{I\cup J} \rightarrow \hat{C}_J \hat{C}_I \; , \\ &[\widetilde{I,J}]^+_\Sigma \quad \text{sign after merging } \hat{C}^\dagger_I \hat{C}^\dagger_J \rightarrow \hat{C}^\dagger_{I\cup J} \; , \\ &[\widetilde{I,J}]^-_\Sigma \quad \text{sign after merging } \hat{C}_I \hat{C}_J \rightarrow \hat{C}_{I\cup J} \; . \end{aligned} \tag{2.1.4}$$

All operators are assumed to be normally ordered before and after the process. The reversed ordering of the annihilation and creation operators ensures that

$$[\overrightarrow{I,J}]^+_\Sigma = [\overleftarrow{J,I}]^-_\Sigma = [\widetilde{I\backslash J, J}]^+_\Sigma = [\widetilde{J, I\backslash J}]^-_\Sigma \; . \tag{2.1.5}$$

Many of the following relations include sums over all subsets of an index set. The summation over all subsets always includes the summation over the empty set if not stated otherwise,

$$\sum_{J\subset I} \ldots = \sum_{J\subset I \vee J=\varnothing} \ldots \ . \tag{2.1.6}$$

Moreover, a prime on a sum which runs over two index sets indicates that both sets must have the same cardinality $|I|$ of a set $I \subset \Omega$,

$$\sum_{I_1,I_2}{}' = \sum_{\substack{I_1,I_2 \\ |I_1|=|I_2|}} \ . \tag{2.1.7}$$

2.1.3 Gutzwiller correlator

The Gutzwiller correlator is given by the product of the local Gutzwiller correlators for all sites l on a lattice Λ.

$$\hat{P}_G = \prod_{l\in\Lambda} \hat{P}_l \ . \tag{2.1.8}$$

A local Gutzwiller operator acts on the site l only. If the context does not lead to any ambiguities, the local index l will be dropped. In this work, we will restrict ourselves to the homogeneous case where the variational parameters in $\hat{P}_l$ are the same for all lattice sites. However, this restriction arises from numerical difficulties only, and most of the derivations can also be performed with an inhomogeneous set of Gutzwiller parameters.

The local Gutzwiller operator is given by

$$\hat{P}_i = \sum_{I_1,I_2} \lambda_{I_1,I_2} \left(|I_1\rangle\langle I_2|\right)_i \ , \tag{2.1.9}$$

$$\hat{P}_i^\dagger \hat{P}_i \ = \sum_{I_1,I_2} \overline{\lambda}_{I_3,I_4}\lambda_{I_1,I_2} \left(|I_1\rangle\langle I_2|I_4\rangle\langle I_3|\right)_i = \sum_{I_1,I_2} \tilde{\lambda}_{I_1,I_2} \left(|I_1\rangle\langle I_2|\right)_i \ , \tag{2.1.10}$$

with $\tilde{\lambda}_{I_1,I_2} = \sum_J \overline{\lambda}_{I_2,J}\lambda_{I_1,J}$. The bare denotes the complex conjugation. The $2^{n_B} \times 2^{n_B}$ complex coefficients parameterize the local Gutzwiller correlator. In chapter 3, we can safely assume that the parameters are real and further constraints such as particle number conservation apply.

The local projector which acts on the site i can be written explicitly as

$$\left(|I_1\rangle\langle I_2|\right)_i \simeq \prod_{l\in\Lambda\setminus i} \mathrm{Id}_l \otimes \left(|I_1\rangle\langle I_2|\right)_i \ , \tag{2.1.11}$$

where Id_l represents the identity operator on site l. Any fermionic projector can be transformed into an expression containing ladder operators only. For example,

$$|I_1\rangle\langle I_2| = \hat{C}_{I_1}^\dagger \hat{C}_{I_2} \prod_{\sigma\in\overline{I_1\cup I_2}} (1-\hat{n}_\sigma) \ . \tag{2.1.12}$$

A lot of the algebraic transformations in this work require a redefinition of index sets. As an example, we express the square of the local Gutzwiller operator in terms of ladder operators,

$$\begin{aligned} P^\dagger P &= \sum_{I_1,I_2} \tilde{\lambda}_{I_1,I_2} |I_1\rangle\langle I_2| \\ &= \sum_{I_1,I_2} \tilde{\lambda}_{I_1,I_2} \hat{C}^\dagger_{I_1} \hat{C}_{I_2} \prod_{\sigma \in \overline{I_1 \cup I_2}} (1-\hat{n}_\sigma) \\ &= \sum_{I_1,I_2} \tilde{\lambda}_{I_1,I_2} \sum_{J \subset \overline{I_1 \cup I_2}} (-1)^{|J|} \hat{C}^\dagger_{I_1} \hat{C}_{I_2} \hat{C}^\dagger_J \hat{C}_J \\ &= \sum_{I_1,I_2} \tilde{\lambda}_{I_1,I_2} \sum_{J \subset \overline{I_1 \cup I_2}} (-1)^{|J|} \widetilde{[I_1, J]}^+_\Sigma \widetilde{[J, I_2]}^-_\Sigma \hat{C}^\dagger_{I_1 \cup J} \hat{C}_{I_2 \cup J} \\ &= \sum_{I_1,I_2} \sum_{J \subset I_1 \cup I_2} \tilde{\lambda}_{I_1 \setminus J, I_2 \setminus J} (-1)^{|J|} \overrightarrow{[I_1, J]}^+_\Sigma \overleftarrow{[J, I_2]}^-_\Sigma \hat{C}^\dagger_{I_1} \hat{C}_{I_2} \\ &= \sum_{I_1,I_2} Y_{I_1,I_2} \hat{C}^\dagger_{I_1} \hat{C}_{I_2} \ . \end{aligned} \tag{2.1.13}$$

In this way we can define the coefficients Y_{I_1,I_2} as

$$Y_{I_1,I_2} = \sum_{J \subset I_1 \cup I_2} \tilde{\lambda}_{I_1 \setminus J, I_2 \setminus J} (-1)^{|J|} \overrightarrow{[I_1, J]}^+_\Sigma \overleftarrow{[J, I_2]}^-_\Sigma \ . \tag{2.1.14}$$

2.1.4 Gutzwiller wave function

A single particle product state (SPPS) can always be cast in the form

$$|\Psi_0\rangle = \prod_{k,\gamma} \hat{h}^\dagger_{k,\gamma} |\mathrm{vac}\rangle \tag{2.1.15}$$

in some fermionic basis

$$\hat{h}^\dagger_{k,\gamma} = \sum_{k,\gamma} U^{i,k}_{\sigma,\gamma} \ \hat{c}^\dagger_{i,\sigma} \ . \tag{2.1.16}$$

We will assume that the SPPS are normalized, $\langle \Psi_0 | \Psi_0 \rangle = 1$, and that the canonical commutation relations $\{\hat{h}^\dagger_{k,\gamma}, \hat{h}_{k',\gamma'}\} = \delta_{kk'}\delta_{\gamma,\gamma'}$ hold. Now, we define the Gutzwiller wave function as

$$|\Psi_G\rangle = \hat{P}_G |\Psi_0\rangle \ . \tag{2.1.17}$$

In the remaining part of this work, we optimize the Gutzwiller correlator $\hat{P}_G$ and the SPPS $|\Psi_0\rangle$ so that the approximate ground state energy

$$E_G = \langle \hat{H} \rangle_G = \frac{\langle \Psi_G | \hat{H} | \Psi_G \rangle}{\langle \Psi_G | \Psi_G \rangle} \tag{2.1.18}$$

becomes minimal. The evaluation of expectation values with the Gutzwiller wave function poses a very difficult many-body problem.

2.1.5 Density matrix and Wick's theorem

The single-particle density matrix of a state $|\Psi\rangle$ is defined as

$$\rho_{ij}^{\sigma\sigma'} = \langle\Psi|\hat{c}_{i\sigma}^{\dagger}\hat{c}_{j\sigma'}|\Psi\rangle \,. \tag{2.1.19}$$

In the following, we will simply call this quantity 'density matrix'. The local density matrix is given by

$$\rho_{\sigma\sigma'} = \rho_{ii}^{\sigma\sigma'} \,. \tag{2.1.20}$$

The diagonal entries of the density matrix and the averaged density are given by $n_\sigma = \langle\hat{n}_\sigma\rangle$ and $n_0 = 1/(2n_B)\sum_\sigma n_\sigma$, respectively. The density matrix of a single-particle product state $|\Psi_0\rangle$ is given by

$$\begin{aligned}\rho_{ij}^{\sigma\sigma'} &= \langle\Psi_0|\hat{c}_{i\sigma}^{\dagger}\hat{c}_{j\sigma'}|\Psi_0\rangle = \sum_{k,k',\gamma,\gamma'} U_{\gamma\sigma}^{ki}\bar{U}_{\gamma'\sigma'}^{k'j}\langle\Psi_0|\hat{h}_{k\gamma}^{\dagger}\hat{h}_{k'\gamma'}|\Psi_0\rangle \\ &= \sum_{k,\gamma}^{\text{occ.}} U_{\gamma\sigma}^{ki}\bar{U}_{\gamma\sigma'}^{kj} \,,\end{aligned} \tag{2.1.21}$$

where the sum over (k,γ) in the last line runs only over all occupied configurations (k,γ).

Wick's theorem states that the expectation value of an operator will be given by the sum over all possible contractions of the operator if $|\Psi_0\rangle$ is a single-particle product state [15]. For example,

$$\begin{aligned}\langle\Psi_0|\hat{c}_{l\sigma_1}^{\dagger}\hat{c}_{k\sigma_2}^{\dagger}\hat{c}_{k\sigma_3}\hat{c}_{l\sigma_4}|\Psi_0\rangle &= \{\hat{c}_{l\sigma_1}^{\dagger}\hat{c}_{k\sigma_2}^{\dagger}\hat{c}_{k\sigma_3}\hat{c}_{l\sigma_4}\}_{\rho_0} \\ &= \{\dot{\hat{c}}_{l\sigma_1}^{\dagger}\ddot{\hat{c}}_{k\sigma_2}^{\dagger}\ddot{\hat{c}}_{k\sigma_3}\dot{\hat{c}}_{l\sigma_4}\}_{\rho_0} + \{\dot{\hat{c}}_{l\sigma_1}^{\dagger}\ddot{\hat{c}}_{k\sigma_2}^{\dagger}\dot{\hat{c}}_{k\sigma_3}\ddot{\hat{c}}_{l\sigma_4}\}_{\rho_0} \\ &= \rho_{ll}^{\sigma_1\sigma_4}\rho_{kk}^{\sigma_2\sigma_3} - \rho_{lk}^{\sigma_1\sigma_3}\rho_{kl}^{\sigma_2\sigma_4} \,,\end{aligned} \tag{2.1.22}$$

where the markers over the operators indicate which operators are contracted with each other. Furthermore, we define the symbol $\{\ldots\}_{\rho_0}$ as the sum over all contraction of the operators inside the braces. We will drop the density matrix in the subindex when it does not lead to ambiguities. In accordance with the above definitions, we set $\{\hat{C}_{\varnothing}^{\dagger}\hat{C}_{\varnothing}\} := 1$. Diagrams can be used to write down all contractions in a symbolic way. For example, the contractions in Eq. (2.1.22) are shown in Fig. (2.5). Furthermore, we can express all local contractions through the determinant of the matrix ρ_{I_1,I_2} which is the submatrix of the elements ρ_{σ_1,σ_2} of the local density matrix with $\sigma \in I_1$ and $\sigma' \in I_2$.

$$\{\hat{C}_{I_1}^{\dagger}\hat{C}_{I_2}\} = \det[\rho_{I_1,I_2}] \,.$$

Note that the indices must be in normal order.

2.1.6 Single-particle transformations

A single-particle transformation

$$\hat{c}_\sigma = U_{\sigma,\sigma'}\hat{c}_{\sigma'} \tag{2.1.23}$$

$$\langle\Psi_0|\hat{c}^\dagger_{l\sigma_1}\hat{c}^\dagger_{k\sigma_2}\hat{c}_{k\sigma_3}\hat{c}_{l\sigma_4}|\Psi_0\rangle =$$

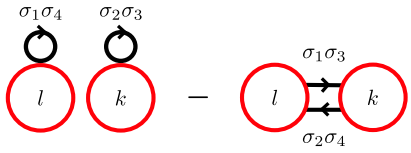


Figure 2.5: All possible contractions of the operators on the left hand side can be symbolized as diagrams. A line symbolizes the corresponding entry of the density matrix. Lines that start and end at the same node are called local loops or Hartree bubbles. The arrow points to the creation operator. All operators that have the same lattice index belong to the same node which is depicted as a red circle. The two contributions have a different topology. The nodes of the left diagram are disconnected while the nodes of the right diagram are connected. The lattice and site index as well as the arrows are dropped if only the structure of the diagram is of importance.

leads to a transformation of the multiparticle states

$$|I'\rangle = \hat{C}^\dagger_{I'}|\mathrm{vac}\rangle = \sum_{I,|I|=|I'|} \det[\bar{U}_{I,I'}]|I\rangle = \sum_I \bar{W}_{I,I'}|I\rangle\,, \tag{2.1.24}$$

where $U_{I,I'}$ is the submatrix of the elements $U_{\sigma,\sigma'}$ with column index $\sigma \in I$ and row index $\sigma' \in I'$. The indices must be normally ordered. The elements of $W_{I,I'}$ are only nonvanishing if $|I| = |I'|$. This transformation will be particularly useful when we compute single-particle transformations induced by the Gutzwiller correlator. This embedding will be denoted ι in the following

$$\iota : U_{\sigma,\sigma'} \to W_{I,I'}\,. \tag{2.1.25}$$

2.2 Hartree-Fock operators

The exclusion of local loops is the first step in our diagrammatic expansion of expectation values for Gutzwiller wave functions. This means that none of the nodes in our diagrams are connected to themselves. In order to achieve this goal, we define Hartree-Fock (HF) operators so that all terms with internal contractions cancel out.

As a first step, we introduce the HF-operators and give explicit expressions. Then, we construct an expansion of an arbitrary operator in terms of HF-operators. We show the consistency of these mappings afterwards. As a further step, we transform the square of the Gutzwiller operator into the HF-operator representation. Finally, we consider external nodes and map them onto HF-operators.

2.2.1 Definition

In this section, a closed expression of the HF-operators is derived for any given set of fermionic creation/annihilation operators $\hat{a}_i$. Consider a product of ladder operators $\hat{a}_1\,\hat{a}_2\ldots\hat{a}_n$. For the corresponding HF-operator $(\hat{a}_1\,\hat{a}_2\ldots\hat{a}_n)^{\mathrm{HF}}$, the evaluation of

$$\{\hat{O}\,(\hat{a}_1\,\hat{a}_2\ldots\hat{a}_n)^{\mathrm{HF}}\} \tag{2.2.1}$$

shall, by definition, not include contractions between any pairs of operators $\hat{a}_1, \ldots, \hat{a}_n$. In the following, we use the notation

$$\{\hat{a}_1\ \hat{a}_2 \ldots \hat{a}_n\}_m \tag{2.2.2}$$

where m denotes the number of internal contractions, e.g., for $m = 1$,

$$\{\hat{a}_1\ \hat{a}_2 \ldots \hat{a}_n\}_1 = \sum_{j<k} (-1)^{j+k+1} \{\hat{a}_j \hat{a}_k\} \hat{a}_1 \ldots \hat{a}_{j-1} \hat{a}_{j+1} \ldots\ldots \hat{a}_{k-1} \hat{a}_{k+1} \ldots \hat{a}_n. \tag{2.2.3}$$

Each internal contraction reduces the number of operators by two.

As a first step, we write down the HF-operator as a sum of the operator chain and some additional terms

$$(\hat{a}_1\ \hat{a}_2 \ldots \hat{a}_n)^{\,\mathrm{HF}} = (\hat{a}_1\ \hat{a}_2 \ldots \hat{a}_n) + \ldots\ . \tag{2.2.4}$$

The sum over all contractions of the operator $\hat{O}$ and the right-hand side of Eq. (2.2.4) generates the desired term in which all of the operators $\hat{a}_i$ are contracted with $\hat{O}$. Additionally, we get the sum over all terms in which all but two (four, six, etc.) operators are contacted with $\hat{O}$. Fig. 2.6 shows the sum over all of these terms which are sorted by the number of contractions with $\hat{O}$.

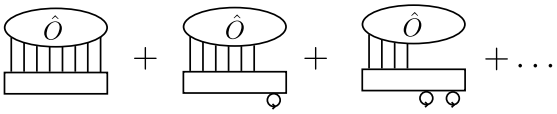
$\{\hat{O}\ \hat{a}_1 \hat{a}_2 \ldots \hat{a}_n\} =$  $+ \ldots$

Figure 2.6: Contraction of an operator product $\hat{a}_1 \ldots \hat{a}_n$ with an arbitrary operator $\hat{O}$, expressed as a sum over terms with the same number of operators performing external contractions. The rectangle represents the operator chain and the ellipse represents the operator $\hat{O}$. Only contractions involving operators $\hat{a}_i$ are shown as solid lines. A self-closing line depicts an internal contraction. Each internal contraction reduces the number of operators by two.

In order to get rid of the extra terms, we subtract them from the original operator chain in an iterative procedure. We start with the subtraction of all terms with one internal contraction and redefine the HF-operator as

$$(\hat{a}_1\ \hat{a}_2 \ldots \hat{a}_n)^{\,\mathrm{HF}} = (\hat{a}_1\ \hat{a}_2 \ldots \hat{a}_n) - \{a_1\ a_2 \ldots a_n\}_1 + \ldots\ . \tag{2.2.5}$$

A further internal contraction of the new term leads to a term with two internal contractions again. This time, two different terms with one internal contraction lead to the same term with two internal contractions. This is why we must multiply the new terms with two or more contractions with the weight $\binom{m}{1}$, where m gives the number of internal contractions. This is shown symbolically in Fig. 2.7.

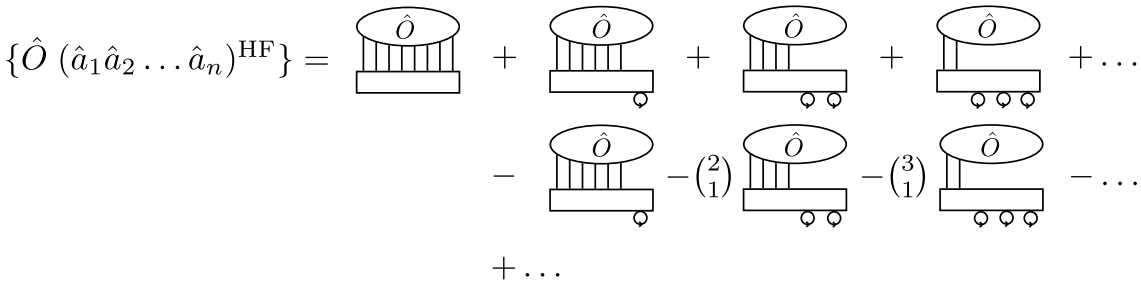


Figure 2.7: Summary of all terms after contracting all terms in Eq. (2.2.5) with $\hat{O}$.

In the next iteration step, we have to take care of the remaining terms with two internal contractions. The total weight of these terms is negative, so that we have to add them to the HF-operators in Eq. 2.2.5

$$(\hat{a}_1\ \hat{a}_2\ldots\hat{a}_n)^{\rm HF} = (\hat{a}_1\ \hat{a}_2\ldots\hat{a}_n) - \{a_1\ a_2\ldots a_n\}_1 + \{a_1\ a_2\ldots a_n\}_2 + \ldots\ . \quad (2.2.6)$$

Inserting Eq. (2.2.6) into Eq. (2.2.1) leads to terms with none or at least three internal contractions. One can easily see that the sum over all terms with three internal contraction is given by $-\binom{3}{0}+\binom{3}{1}-\binom{3}{2}=1$ so that these terms have to be subtracted from our definition of the HF-operators as a next step. A summary of this procedure is shown in Fig. 2.8.

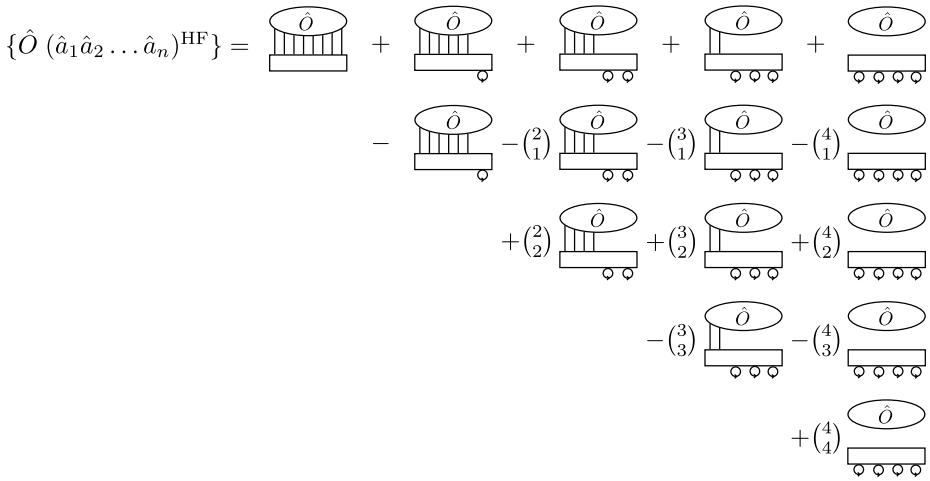


Figure 2.8: Summary of all terms obtained after four recursive steps for $n = 8$.

Let us assume that $n > 2m$ and that we have already added up all terms with at most m internal contractions multiplied with an alternating sign,

$$(\hat{a}_1\ \hat{a}_2\ldots\hat{a}_n)^{\rm HF} = \sum_{k=0}^{m}(-1)^k\{a_1\ a_2\ldots a_n\}_k\ . \quad (2.2.7)$$

In the next step we have to get rid of the

$$\sum_{k=0}^{m}(-1)^k\binom{m+1}{k} = (-1)^m \quad (2.2.8)$$

terms with $m+1$ internal contraction. Thus we can simply add the next term to our alternating sum until we end up with all operators contracted with each other, or a single operators $\hat{a}_i$ left to be contracted with $\hat{O}$.

The recursive procedure stops here and we can give a closed expression for the HF-operator of an arbitrary operator

$$(\hat{a}_1\ \hat{a}_2 \dots \hat{a}_n)^{\mathrm{HF}} = \sum_{k=0}^{[n]} (-1)^k \{a_1\ a_2 \dots a_n\}_k\ , \tag{2.2.9}$$

where $[n]$ denotes the next smallest even number. The mapping is linear in its argument and any operator can be written as a sum of ladder operators so that Eq. (2.2.9) serves as our general definition. Note that a closed expressions for the definition of the HF-operators can be found in [14], although without derivation.

As a concluding step, the sum over the number of contractions in Eq. (2.2.9) can be translated to a sum over the index sets of the operators that are contracted. We find

$$(A_I)^{\mathrm{HF}} = \sum_{\substack{J\subset I \\ |J|\text{even}}} (-1)^{|J|/2} \{A_I\}_J\ , \tag{2.2.10}$$

where the index J of the curly braces denotes the subset of I which are to be contracted. We split up the operator $\hat{A}$ in terms of creation and annihilation operators

$$\begin{aligned}
\left(\hat{C}^\dagger_{I_1}\hat{C}_{I_2}\right)^{\mathrm{HF}} &= {\sum_{\substack{J_1\subset I_1,\\ J_2\subset I_1}}}' (-1)^{|J_1|} \{\hat{C}^\dagger_{I_1}\hat{C}_{I_2}\}_{J_1,J_2} \\
&= {\sum_{\substack{J_1\subset I_1,\\ J_2\subset I_1}}}' (-1)^{|J_1|} [\overrightarrow{I_1,J_1}]^+_\Sigma [\overleftarrow{J_2,I_2}]^-_\Sigma \{\hat{C}^\dagger_{J_1}\hat{C}_{J_2}\} \hat{C}^\dagger_{I_1\setminus J_1}\hat{C}_{I_2\setminus J_1} \\
&= {\sum_{\substack{J_1\subset I_1,\\ J_2\subset I_1}}}' (-1)^{|I_1\setminus J_1|} [\overleftarrow{J_1,I_1}]^+_\Sigma [\overrightarrow{I_2,J_2}]^-_\Sigma \{\hat{C}^\dagger_{I_1\setminus J_1}\hat{C}_{I_2\setminus J_2}\} \hat{C}^\dagger_{J_1}\hat{C}_{J_2}\ ,
\end{aligned} \tag{2.2.11}$$

where the index sets $J_1 \subset I_1$ and $J_2 \subset I_2$ denote the creation and annihilation operators to be contracted with each other in the first line. In the second and third line, all operators inside the braces are contracted with each other. In order to separate the contracted operators, the remaining operators have to be pulled out of the operators $\hat{C}^\dagger_{I_1}$ and $\hat{C}_{I_2}$ which generates additional sign factors. We summarize the result as

$$\left(\hat{C}^\dagger_{I_1}\hat{C}_{I_2}\right)^{\mathrm{HF}} = \sum_{J_1,J_2} \chi^{I_1,I_2}_{J_1,J_2}\ \hat{C}^\dagger_{J_1}\hat{C}_{J_2}\ , \tag{2.2.12}$$

with

$$\chi^{I_1I_2}_{J_1,J_2} = {\sum_{\substack{J_1\subset I_1,\\ J_2\subset I_1}}}' (-1)^{|I_1\setminus J_1|} [\overleftarrow{J_1,I_1}]^+_\Sigma [\overrightarrow{I_2,J_2}]^-_\Sigma \{\hat{C}^\dagger_{I_1\setminus J_1}\hat{C}_{I_2\setminus J_2}\}\ . \tag{2.2.13}$$

2.2.2 Expansion of operators

In our diagrammatic expansion, we must transform all operators $\hat{C}^\dagger_{K_1}\hat{C}_{K_2}$ in terms of HF-operators,

$$\hat{C}^\dagger_{K_1}\hat{C}_{K_2} = \sum_{I_1,I_2} X^{K_1,K_2}_{I_1,I_2} \left(\hat{C}^\dagger_{I_1}\hat{C}_{I_2}\right)^{\mathrm{HF}} . \tag{2.2.14}$$

Let us contract both sides with an arbitrary operator $\hat{O}$. In order to determine the value of the coefficient $X^{K_1K_2}_{I_1,I_2}$, we have to evaluate the term where the operators $\hat{C}^\dagger_{I_1}$ and $\hat{C}_{I_2}$ with $I_1 \subset K_1$, $I_2 \subset K_2$ form the external contractions with $\hat{O}$ which is shown in Fig. 2.9.

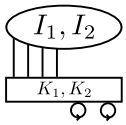
$X^{K_1K_2}_{I_1,I_2} =$ 

Figure 2.9: Coefficient $X^{K_1K_2}_{I_1I_2}$ corresponding to the term in which the operators $\hat{C}^\dagger_{I_1}$ and $\hat{C}_{I_2}$ with $I_1 \subset K_1$, $I_2 \subset K_2$ form external contractions with $\hat{O}$. The ellipse represents the operator $\hat{O}$ while the rectangle represents the operator $\hat{C}^\dagger_{K_1}\hat{C}_{K_2}$.

We need to shift the operators that are reserved for the external contractions to the front of the operator $\hat{O}$, and contract all remaining operators internally. The operator $\hat{O}$ can be chosen without any restrictions and just indicates which operators are reserved for an external contraction. The contraction with $\hat{O}$ can be carried out symbolically by a replacement of $\hat{C}^\dagger_{I_1}\hat{C}_{I_2}$ with the HF-operators $(\hat{C}^\dagger_{I_1}\hat{C}_{I_2})^{\mathrm{HF}}$. Thus, we get

$$X^{K_1,K_2}_{I_1,I_2} = [\overleftarrow{I_1,K_1}]^+_\Sigma [\overrightarrow{K_2,I_2}]^-_\Sigma \{\hat{C}^\dagger_{K_1\setminus I_1}\hat{C}_{K_2\setminus I_2}\} . \tag{2.2.15}$$

We can insert this result into Eq. (2.2.14) to find

$$\hat{C}^\dagger_{K_1}\hat{C}_{K_2} = \sum_{I_1,I_2} [\overleftarrow{I_1,K_1}]^+_\Sigma [\overrightarrow{K_2,I_2}]^-_\Sigma \{\hat{C}^\dagger_{K_1\setminus I_1}\hat{C}_{K_2\setminus I_2}\} \left(\hat{C}^\dagger_{I_1}\hat{C}_{I_2}\right)^{\mathrm{HF}} . \tag{2.2.16}$$

The transformation of an operator to its HF-representation depends on the local density matrix but it is a unique and invertible mapping. The density matrix itself does not have to fulfill any requirements and can be chosen arbitrarily. A straightforward proof of the consistency of Eq. (2.2.16) and Eq. (2.2.11) is given in appendix A.1.

2.2.3 Transformation of the Gutzwiller correlator

As a next step we transform the square of the local Gutzwiller operator into a sum of HF-operators,

$$P^\dagger P = \sum_{I_1,I_2} Y_{I_1,I_2}\hat{C}^\dagger_{I_1}\hat{C}_{I_2} = \sum_{I_1,I_2} X_{I_1,I_2} \left(\hat{C}^\dagger_{I_1}\hat{C}_{I_2}\right)^{\mathrm{HF}} . \tag{2.2.17}$$

The mapping between the normal operators $\hat{C}^\dagger_{I_1}\hat{C}_{I_2}$ and their HF-operators $(\hat{C}^\dagger_{I_1}\hat{C}_{I_2})^{\mathrm{HF}}$ is linear. Therefore, we transform all operators and summarize terms with equal HF-operators as shown in Fig. 2.10.

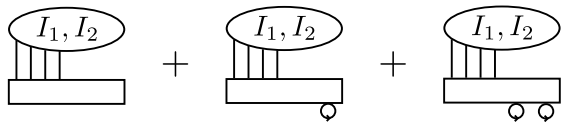
$X_{I_1,I_2} =$  $+\ldots$

Figure 2.10: X_{I_1,I_2} sums up all terms in which the external contributions are formed by the operators $\hat{C}^\dagger_{I_1}\hat{C}_{I_2}$.

As seen from the figure, the coefficient X_{I_1,I_2} is given by the sum over all coefficients Y_{J_1,J_2} with $I_1 \subset J_1$ and $I_2 \subset J_2$ where all remaining operators have been contracted internally. Formally,

$$\begin{aligned} X_{I_1,I_2} &= \sum_{J_1, I_1\subset J_1} \sum_{\substack{J_2, I_2\subset J_2 \\ \wedge |J_1\setminus I_1| = |J_2\setminus I_2|}} Y_{J_1,J_2}\{\hat{C}^\dagger_{J_1\setminus I_1}\hat{C}_{J_2\setminus I_2}\}[\overleftarrow{I_1, J_1}]^+_\Sigma[\overrightarrow{J_2, I_2}]^-_\Sigma \qquad (2.2.18)\\ &= {\sum_{J_1\subset \bar{I}_1, J_2\subset \bar{I}_2}}' Y_{J_1\cup I_1, J_2\cup I_2}\{\hat{C}^\dagger_{J_1}\hat{C}_{J_2}\}[\widetilde{I_1, J_1}]^+_\Sigma[\widetilde{J_2, I_2}]^-_\Sigma \\ &= {\sum_{J_1\subset \bar{I}_1, J_2\subset \bar{I}_2}}' \sum_{\substack{J_3\subset \\ (I_1\cup J_1)\cap(I_2\cup J_2)}} \tilde{\lambda}^{(I_2\cup J_2)\setminus J_3}_{(I_1\cup J_1)\setminus J_3}\{\hat{C}^\dagger_{J_1}\hat{C}_{J_2}\} \\ &\quad [\overrightarrow{I_1\cup J_1, J_3}]^+_\Sigma\,[\overleftarrow{J_3, I_2\cup J_2}]^-_\Sigma[\widetilde{I_1, J_1}]^+_\Sigma\,[\widetilde{J_2, I_2}]^-_\Sigma\,. \end{aligned}$$

Eq. (2.2.18) permits the calculation of the coefficients X_{I_1,I_2} for given variational parameters λ_{I_1,I_2}. The reverse mapping can be derived by a projection of Eq. (2.2.18) to the states $\langle I_1|$ and $|I_2\rangle$

$$\begin{aligned} P^\dagger P &= \sum_{I_1,I_2} \tilde{\lambda}_{I_1,I_2}|I_1\rangle\langle I_2| \qquad (2.2.19)\\ &= \sum_{I_1,I_2} X_{I_1,I_2}\left(\hat{C}^\dagger_{I_1}\hat{C}_{I_2}\right)^{\mathrm{HF}} \\ &= {\sum_{I_1,I_2}}' \sum_{\substack{J_1\subset \bar{I}_1 \\ J_2\subset \bar{I}_2}} X_{I_1\cup J_1, I_2\cup J_2}(-1)^{|J_1|}[\widetilde{I_1, J_1}]^+_\Sigma[\widetilde{J_2, I_2}]^-_\Sigma\{\hat{C}^\dagger_{I_1}\hat{C}_{I_2}\}\hat{C}^\dagger_{J_1}\hat{C}_{J_2} \end{aligned}$$

so that

$$\tilde{\lambda}_{I_1,I_2} = {\sum_{I_1,I_2}}' \sum_{\substack{J_1\subset \bar{I}_1 \\ J_2\subset \bar{I}_2}} X_{I_1\cup J_1, I_2\cup J_2}(-1)^{|J_1|}[\widetilde{I_1, J_1}]^+_\Sigma[\widetilde{J_2, I_2}]^-_\Sigma\{\hat{C}^\dagger_{I_1}\hat{C}_{I_2}\}\langle I_1|\hat{C}^\dagger_{J_1}\hat{C}_{J_2}|I_2\rangle\,. \qquad (2.2.20)$$

In this way, the variational parameters $\tilde{\lambda}_{I_1,I_2}$ can be used as functions of the coefficients X_{I_1,I_2}.

One might think to use the coefficients X_{I_1,I_2} as free variational parameters. However, there are some obstacles. Most importantly, the coefficients λ_{I_1,I_2} and not the coefficients $\tilde{\lambda}_{I_1,I_2}$ enter the external nodes of the diagrammatic analysis which will be derived in the next subsection. Therefore, the coefficients X_{I_1,I_2} would have to be decomposed numerically into the coefficients λ_{I_1,I_2}. We would not be able to obtain any analytical expressions for the derivation of the internal nodes with respect to the variational parameters. Any constraints on the coefficients $\tilde{\lambda}_{I_1,I_2}$ would also have to be incorporated into the coefficients X_{I_1,I_2}. For example one might think of a non-Hermitian set of coefficients λ_{I_1,I_2}. For these reasons we keep the coefficients λ_{I_1,I_2} as our variational parameters.

2.2.4 Transformation of external operators

As will be shown in section 2.3, the evaluation of expectation values of any operator $\hat{O}$ with respect to the Gutzwiller wave function $|\Psi_G\rangle$ requires the transformation of the sandwich $\hat{Q} = \hat{P}^\dagger \hat{O} \hat{P}$ into the HF operator representation.

As a first step, the sandwich is expressed as

$$\hat{Q}(\hat{O}) = \hat{P}^\dagger \hat{O} \hat{P} = \sum_{I_1,I_2} K_{I_1,I_2}(\hat{O}) |I_1\rangle\langle I_2| \tag{2.2.21}$$

with

$$K_{I_1,I_2}(\hat{O}) = K^{I_2}_{I_1}(\hat{O}) = \sum_{I_3,I_4} \bar{\lambda}_{I_3,I_1} \lambda_{I_4,I_2} \langle I_3|\hat{O}|I_4\rangle \,. \tag{2.2.22}$$

Then, we can use the HF mapping to find

$$\hat{Q}(\hat{O}) = \sum_{I_1,I_2} Q_{I_1,I_2}(\hat{O}) \left(\hat{C}^\dagger_{I_1} \hat{C}_{I_2}\right)^{\mathrm{HF}} \,, \tag{2.2.23}$$

which still depends on the operator $\hat{O}$. Here,

$$Q_{I_1,I_2}(\hat{O}) = \sum_{\substack{J_1 \subset \bar{I}_1 \\ J_2 \subset \bar{I}_2}}{}' \sum_{\substack{J_3 \subset \\ (I_1 \cup J_1) \cap (I_2 \cup J_2)}} K^{(I_2 \cup J_2)\setminus J_3}_{(I_1 \cup J_1)\setminus J_3}(\hat{O}) (-1)^{|J_3|} \{\hat{C}^\dagger_{J_1} \hat{C}_{J_2}\} \\ \overrightarrow{[I_1 \cup J_1, J_3]}^+_\Sigma [\overleftarrow{J_3, I_2 \cup J_2}]^-_\Sigma \widetilde{[I_1, J_1]}^+_\Sigma \widetilde{[J_2, I_2]}^-_\Sigma \,, \tag{2.2.24}$$

where the coefficients K_{I_1,I_2} play the role of the coefficients $\tilde{\lambda}_{I_1,I_2}$.

2.3 Linked cluster theorem

In this section we derive a diagrammatic expansion of the Gutzwiller wave function on a finite-dimensional lattice. We compute expectation values $\langle \hat{O} \rangle_G$ in terms of a series expansion where the nth order is defined by all connected diagrams with n

internal vertices. These diagrams have to be evaluated numerically afterwards. The derivation is given for an arbitrary local configuration space and does not assume any symmetries such as particle number conservation ($|I_1| = |I_2|$) or a homogeneous set of Gutzwiller operators. The second part of this section reviews previous works on this topic. The evaluation of the diagrammatic expansion on an infinite dimensional lattice is discussed in the third subsection.

2.3.1 Derivation

We consider the expectation value of a local operator $\hat{O}_i$

$$\langle \hat{O}_i \rangle_G = \frac{\langle \Psi_G | \hat{O}_i | \Psi_G \rangle}{\langle \Psi_G | \Psi_G \rangle} \tag{2.3.1}$$

with

$$|\Psi_G\rangle = \prod_l \hat{P}_l |\Psi_0\rangle \,, \tag{2.3.2}$$

and $|\Psi_0\rangle$ is an arbitrary single-particle product state (SPPS). The same calculations can be performed for nonlocal operators like $\langle \hat{c}_i^\dagger \hat{c}_f \rangle_G$. Sums and products run over all lattice sites $l \in \Lambda$. Note that we are interested in the thermodynamic limit, $L_\Lambda \to \infty$. However, it turns out that the following analysis can also be performed on finite lattices.

As a first step, we follow the analysis for the single-band case derived in [8] and partly worked out for the multi-band case in infinite dimensions in [14]. In the numerator of Eq. (2.3.1) we pull the Gutzwiller correlators with indices $l \neq i$ to the right side of $\hat{O}_i$ and denote the sandwich $\hat{P}_i^\dagger \hat{O}_i \hat{P}_i$ as $\hat{Q}_i$,

$$\langle \Psi_G | \hat{O}_i | \Psi_G \rangle = \langle \Psi_0 | \hat{P}_i^\dagger \hat{O}_i \hat{P}_i \prod_{l \in \Lambda \setminus i} \hat{P}_l^\dagger \hat{P}_l \, | \Psi_0 \rangle = \langle \Psi_0 | \hat{Q}_i \prod_{l \in \Lambda \setminus i} \hat{P}_l^\dagger \hat{P}_l \, | \Psi_0 \rangle \,. \tag{2.3.3}$$

We apply Wick's theorem and compute all possible contractions

$$\langle \Psi_0 | \hat{Q}_i \prod_{l \in \Lambda \setminus i} \hat{P}_l^\dagger \hat{P}_l \, | \Psi_0 \rangle = \{ \hat{Q}_i \prod_{l \in \Lambda \setminus i} \hat{P}_l^\dagger \hat{P}_l \, \}_\rho \,, \tag{2.3.4}$$

where $\rho_{ij}^{\sigma\sigma'} = \langle \Psi_0 | \hat{c}_{i\sigma}^\dagger \hat{c}_{j\sigma'} | \Psi_0 \rangle$ denotes the density matrix. Proceeding this way, we would keep all local contractions. However, when we employ the analysis in section 2.2, we can transform all operators in such a way that all local contractions vanish. To achieve this goal, we map our operators to HF-operators,

$$\hat{Q} = \sum_{I_1, I_2} Q_{I_1, I_2} \left(\hat{C}_{I_1}^\dagger \hat{C}_{I_2} \right)^{\mathrm{HF}} \,, \tag{2.3.5}$$

$$\hat{P}^\dagger \hat{P} = 1 + \hat{A} = 1 + \sum_{\substack{I_1, I_2 \\ |I_1|, |I_2| > 0}} X_{I_1, I_2} \left(\hat{C}_{I_1}^\dagger \hat{C}_{I_2} \right)^{\mathrm{HF}} \,, \tag{2.3.6}$$

where we set the coefficient $X_{\varnothing,\varnothing} = 1$. This is equal to a rescaling of the Gutzwiller wave function by a scalar factor which is always canceled out by the denominator in Eq. (2.3.1).

All operators in Eq. (2.3.4) are normal ordered because all site indices are different when we apply Wick's theorem. We can set all local entries in ρ to zero because we work with the HF-operators so that all local contractions vanish automatically. Therefore, we can carry out all contractions with a new density matrix

$$\tilde{\rho}_{ij}^{\sigma\sigma'} = \rho_{ij}^{\sigma\sigma'} - \delta_{ij}\, \tilde{\rho}_{ii}^{\sigma\sigma'} \tag{2.3.7}$$

and drop the HF-operator notation at the same time

$$\{\left(\hat{C}_{I_1}^{\dagger}\hat{C}_{I_2}\right)^{\mathrm{HF}} \ldots\}_{\rho_0} \equiv \{\tilde{C}_{I_1}^{\dagger}\tilde{C}_{I_2}\ldots\}_{\tilde{\rho}_0}\ . \tag{2.3.8}$$

Without any nonzero local contraction we get

$$\{\tilde{c}_{is}^{\dagger}\tilde{c}_{is}\ldots\}_{\tilde{\rho}_0} = -\{\tilde{c}_{is}\tilde{c}_{is}^{\dagger}\ldots\}_{\tilde{\rho}_0}\ . \tag{2.3.9}$$

Thus, we can consider our new objects as Graßmann variables instead of fermionic operators. These Graßmann variables are nilpotent

$$\tilde{C}_{I_1}\tilde{C}_{I_2} = 0 \quad \text{if } I_1 \cap I_2 \neq \varnothing\ . \tag{2.3.10}$$

In principle, the introduction of the HF mapping is not a necessary step for the introduction of Graßmann operators as we discuss in the next subsection 2.3.2. All local entries of the new density matrix $\tilde{\rho}$ vanish so that the diagrammatic expansion cannot have nodes with self-closing lines, as shown in Fig. 2.11.

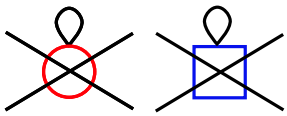


Figure 2.11: All nodes with internal lines cancel because the local entries of $\tilde{\rho}$ are set to zero.

The coefficients Q_{I_1,I_2} and X_{I_1,I_2} are not affected by our mapping so that we can write

$$\tilde{Q}_l = \sum_{I_1,I_2} Q_{I_1,I_2}\tilde{C}_{I_1}^{\dagger}\tilde{C}_{I2} \tag{2.3.11}$$

$$\tilde{A}_l = \sum_{\substack{I_1,I_2 \\ |I_1|,|I_2|>0}} X_{I_1,I_2}\tilde{C}_{I_1}^{\dagger}\tilde{C}_{I2}\ . \tag{2.3.12}$$

The numerator in Eq. (2.3.4) becomes

$$\{\hat{Q}_i \prod_{l\in\Lambda\setminus i}(1+\hat{A}_l)\}_{\rho_0} = \{\tilde{Q}_i \prod_{l\in\Lambda\setminus i}(1+\tilde{A}_l)\}_{\tilde{\rho}_0} \tag{2.3.13}$$

whereas the denominator reads

$$\{\prod_{l\in\Lambda}(1+\hat{A}_l)\}_{\rho_0} = \{\prod_{l\in\Lambda}(1+\tilde{A}_l)\}_{\tilde{\rho}_0} \ . \tag{2.3.14}$$

The lattice site restrictions on the right hand site of Eq. (2.3.13) must be removed before we can apply the linked cluster theorem. Therefore, we define

$$1+\tilde{A}_l = \exp(\tilde{G}_l) \ . \tag{2.3.15}$$

The exponential series expansion stays finite due to the nilpotency of the Graßmann variables. It is crucial that we perform the HF-mapping before we switch to the exponential form of our correlators. The logarithm of Eq. (2.3.15) gives

$$\tilde{G} = \sum_{m>0} \frac{(-1)^{m+1}}{m} \left(\sum_{J_1,J_2} X_{J_1,J_2} \tilde{C}^\dagger_{J_1} \tilde{C}_{J_2} \right)^m \ . \tag{2.3.16}$$

The sum over m is restricted to a finite number of terms again. We can summarize this as

$$\tilde{G} = \sum_{\substack{I_1,I_2 \\ |I_1|,|I_2|>0}} Z_{I_1,I_2} \tilde{C}^\dagger_{I_1} \tilde{C}_{I_2} \ , \tag{2.3.17}$$

with

$$Z_{I_1,I_2} = \sum_{m>0} \frac{(-1)^{m+1}}{m} \sum_{\substack{\{(J_1^s,J_2^s)\} \\ s=1,\ldots,m}} \Sigma[J_1^s,J_2^s] \prod_{s=1}^m X_{J_1^s,J_2^s} \ , \tag{2.3.18}$$

where the sum runs over all (disjunct) partitions $\{(J_1^s,J_2^s)\}$ of the set (I_1,I_2) such that

$$\bigcup_s J_1^s = I_1 \text{ and } \bigcup_s J_2^s = I_2 \ , \tag{2.3.19}$$

and $\Sigma[J_1^s,J_2^s]$ gives the sign which is necessary to convert the operator product $\prod_s \hat{C}^\dagger_{J_1^s}\hat{C}_{J_2^s}$ into normal order again. We apply a similar procedure to the external nodes and define

$$\tilde{Q}_l = \tilde{M}_l \exp(\tilde{G}_l) \ . \tag{2.3.20}$$

We get

$$\tilde{M} = \sum_{I_1,I_2} M_{I_1,I_2} \tilde{C}^\dagger_{I_1} \tilde{C}_{I_2} \ , \tag{2.3.21}$$

with

$$M_{I_1,I_2} = Q_{I_1,I_2} - \sum_{m=1} (-1)^{m+1} \sum_{\substack{\{(J_1^s,J_2^s)\} \\ s=0,\ldots,m}} \Sigma[J_1^s,J_2^s] Q_{J_1^0,J_2^0} \prod_{s=1}^m X_{J_1^s,J_2^s} \ . \tag{2.3.22}$$

These additional redefinitions allows us to rewrite Eq. (2.3.13) as

$$\begin{aligned}\langle\Psi_G|\hat{O}_i|\Psi_G\rangle &= \{\tilde{Q}_i \prod_{l\in\Lambda\setminus i}(1+\tilde{A}_l)\}_{\tilde{\rho}_0} \\ &= \{\tilde{Q}_i \exp(-\tilde{G}_i)\exp(\tilde{G}_i)\prod_{l\in\Lambda\setminus i}\exp(\tilde{G}_l)\}_{\tilde{\rho}_0} \\ &= \{\tilde{M}_i \prod_{l\in\Lambda}\exp(\tilde{G}_l)\}_{\tilde{\rho}_0}\ .\end{aligned} \tag{2.3.23}$$

Note that the site index restriction $l \neq i$ disappeared. Eq. (2.3.14) can be rewritten as

$$\langle\Psi_G|\Psi_G\rangle = \{\prod_{l\in\Lambda}\exp(\tilde{G}_l)\}_{\tilde{\rho}_0}\ . \tag{2.3.24}$$

Now, we are in the position to apply the linked cluster theorem (LCT), as described, e.g., in [15]. Therefore, we pick up all terms with p internal nodes $\tilde{G}_{l_1}\ldots\tilde{G}_{l_p}$. Each operator $\tilde{G}_l$ is built up by a sum of terms but for the moment we consider them as a single node. In order to do this, we can generalize the site index l to a combined index that also labels the coefficients Z_{I_1,I_2}. All $\tilde{G}_l$ commute and the exponential function can be expanded into a product,

$$\begin{aligned}\{\tilde{M}_i\prod_l\exp(\tilde{G}_l)\}_{\tilde{\rho}_0} &= \sum_{p=0}^{L}\sum_{\substack{l_1,\ldots,l_p\\ l_i\neq l_j}}\sum_{n_1,\ldots,n_p}^{n_1+\ldots+n_p=p}\eta_{n_1,\ldots,n_p}\{\tilde{M}_i\tilde{G}_{l_1}^{n_1}\ldots\tilde{G}_{l_p}^{n_p}\}_{\tilde{\rho}_0} \\ &= \sum_{p=0}^{L}\sum_{l_1,\ldots,l_p}\frac{1}{p!}\{\tilde{M}_i\tilde{G}_{l_1}\ldots\tilde{G}_{l_p}\}_{\tilde{\rho}_0}\ ,\end{aligned} \tag{2.3.25}$$

where we abbreviated

$$\eta_{n_1,\ldots,n_p} = \frac{1}{n_{n_1}!\ldots n_{n_p}!}\ . \tag{2.3.26}$$

This shows that we can evaluate the chain of the exponential functions in Eq. (2.3.25) by summing over all index subsets $l_1,\ldots,l_p$.

Let us return to the first line of Eq. (2.3.25). After the contractions have been performed we can split up the internal nodes $\tilde{G}$ into two categories. The first category consists of all nodes that are directly or indirectly linked to the external node $\tilde{M}_i$. This means that some operators of a node $\tilde{G}_{l_i}$ must be contracted with some operators of $\tilde{M}_i$, or with those of a node $\tilde{G}_{l_j}$ that is connected to $\tilde{M}_i$. Here, two nodes will only be connected with each other if they have a common contraction not just because they are placed on the same site. If several nodes lie on the same lattice site, we have to consider all possible combinations to distribute them over the two groups. A careful analysis of the combinatorial prefactors shows that we

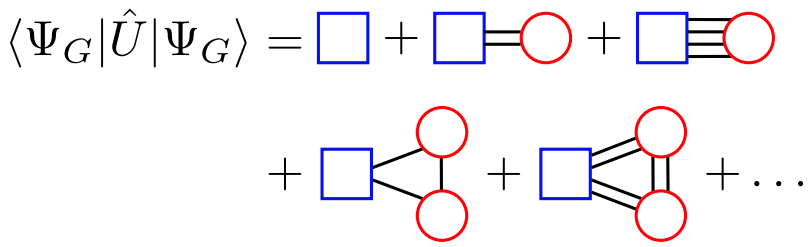


Figure 2.12: The first few connected diagrams that contribute to $\langle \hat{U} \rangle_G$. The blue square represents the external node. The red circles represent the internal nodes. Black lines stand for the single-particle density matrix.

can separate the coefficient $\eta_{n_1,\ldots,n_p}$ into two independent factors for each group.

$$\begin{aligned}\{\tilde{M}_i \prod_l \exp(\tilde{G}_l)\}_{\tilde{\rho}_0} &= \sum_{r,p,q} \delta_{r,q+p} \sum_{\substack{l_1,\ldots,l_p \\ l_i \neq l_j}} \sum_{n_1,\ldots,n_p}^{n_1+\ldots+n_p=p} \eta_{n_1,\ldots,n_p} \{\tilde{M}_i \tilde{G}_{l_1}^{n_1} \ldots \tilde{G}_{l_p}^{n_p}\}_{\tilde{\rho}_0}^{\text{conn.}} \\ &\times \sum_{\substack{k_1,\ldots,k_q \\ k_i \neq k_j}} \sum_{m_1,\ldots,m_q}^{m_1+\ldots+m_q=q} \eta_{m_1,\ldots,m_q} \{\tilde{G}_{k_1}^{m_1} \ldots \tilde{G}_{k_q}^{m_q}\}_{\tilde{\rho}_0} \\ &= \sum_{m=0}^{L} \sum_{l_1,\ldots,l_m} \frac{1}{m!} \{\tilde{M}_i \tilde{G}_{l_1} \ldots \tilde{G}_{l_m}\}^{\text{conn.}} \sum_{q=0}^{L} \sum_{k_1,\ldots,k_q} \frac{1}{q!} \{\tilde{G}_{l_1} \ldots \tilde{G}_{l_q}\}\end{aligned} \tag{2.3.27}$$

The last part of the last line gives the sum over all nodes which are not connected to the external node $\tilde{M}_i$. These unconnected nodes are canceled by the denominator of the expectation value. Finally, we find

$$\frac{\langle \Psi_G | \hat{O}_i | \Psi_G \rangle}{\langle \Psi_G | \Psi_G \rangle} = \sum_{L \subset \Lambda} \frac{1}{|L|!} \{\tilde{M}_i \prod_{l \in L} \tilde{G}_l\}_{\tilde{\rho}_0}^{\text{conn.}}, \tag{2.3.28}$$

where the summation is performed over all subsets L of the lattice Λ. Eq. (2.3.28) constitutes the linked cluster theorem (LCT). The first few diagrams that are needed for the evaluation of the potential energy are shown in Fig. 2.12.

Some of the polynomials of $\tilde{G}_{I_1,I_2,l}$ in Eq. (2.3.23) vanish due to the nilpotency of the Graßmann variables. In contrast to the usual application of the LCT in many body lattice theories, we can apply our expansion for a finite lattice as well. The nilpotency property allows us to add virtually as many nodes as we need to regroup our diagrams in all orders. Note that after the application of the LCT the nodes $\tilde{G}_{I_1,I_2,l}$ are contracted in such a way that all nodes have to be connected to the external nodes M_i. This invalidates the nilpotency of the Graßmann variables inside the curly brackets. Therefore, several nodes can be located on the same site as long as these nodes are only connected indirectly.

2.3.2 Previous approaches

The diagrammatic analysis of the single-band case which includes the HF-operators was first worked out in [8, 9]. There, the correlator $\tilde{A} = x_d \tilde{n}_\uparrow \tilde{n}_\downarrow$ is employed with x_d

being the only nonvanishing X_{I_1,I_2} coefficient. In this case, we can set

$$1 + x_d \tilde{C}^\dagger_{\uparrow,\downarrow} \tilde{C}_{\uparrow,\downarrow} = \exp(x_d \tilde{C}^\dagger_{\uparrow,\downarrow} \tilde{C}_{\uparrow,\downarrow}) \; . \tag{2.3.29}$$

When we are dealing with multiple bands or if we allow terms like $x_\uparrow \tilde{c}^\dagger_\uparrow \tilde{c}_\uparrow$ in our correlator, we need to re-exponentiate our Graßmann operators. This has been overlooked in the deviation of the LCT in the multiband case in [14] although the problem was already noticed in [10].

In principle, the transformation of the ladder operators to HF-operators is not a necessary step for the transformation to Graßman variables. After all operators have been brought to normal ordering every operator in the numerator (denominator) will appear only once. The operators can be mapped to Graßmann variables

$$\hat{c}^\dagger_{i\sigma} \to \tilde{c}^\dagger_{i\sigma} \; , \quad \hat{c}_{i\sigma} \to \tilde{c}_{i\sigma} \quad , \tag{2.3.30}$$

with vanishing anticommutator $\{\tilde{c}^\dagger_{i\sigma}, \tilde{c}_{i\sigma}\} = 0$. In contrast to the Graßmann variables defined in Eq. (2.3.9), the local contractions are still finite. The external and internal nodes can be defined as

$$\begin{aligned} \hat{Q}_l &= \sum_{I_1,I_2} K_{I_1,I_2} \hat{C}^\dagger_{I_1} \hat{C}_{I2} \; , & \tilde{Q}_l &= \sum_{I_1,I_2} K_{I_1,I_2} \tilde{C}^\dagger_{I_1} \tilde{C}_{I2} \; , \\ \hat{A}_l &= \sum_{\substack{I_1,I_2 \\ |I_1|,|I_2|>0}} Y_{I_1,I_2} \hat{C}^\dagger_{I_1} \hat{C}_{I2} \; , & \tilde{A}_l &= \sum_{\substack{I_1,I_2 \\ |I_1|,|I_2|>0}} Y_{I_1,I_2} \tilde{C}^\dagger_{I_1} \tilde{C}_{I2} \; . \end{aligned} \tag{2.3.31}$$

The internal Z_{I_1,I_2} and external M_{I_1,I_2} nodes can be evaluated by a simply replacement of the Q_{I_1,I_2} and X_{I_1,I_2} coefficients with the coefficients Y_{I_1,I_2} and K_{I_1,I_2}. In this approach, the diagrams will also include local lines. That means that we need to include the summation of an arbitrary number of (directly connected) nodes sitting on the same site, in order to sum up all local contractions. For the single-band model the diagrammatic expansion of this case is derived in [4–7], where the cases of one and infinite dimensions are treated analytically. A similar approach for a three-flavor system with an Gutzwiller correlator of the form $1 + \alpha \hat{n}_1 \hat{n}_2 \hat{n}_3$ can be found in [16], where the local contractions are still present. The transformation to an exponential function is again trivial. The big advantage of the diagrammatic setup developed in subsection 2.3.1 lies in the fact that we obtain simple, explicit results in the limit of infinite dimensions.

2.3.3 Limit of infinite dimensions

A scaling analysis of the kinetic energy operator [6, 8, 9] shows that the lines of the density matrix scale with the lattice dimension d as

$$\rho_{i,j} \sim (\sqrt{2d})^{-||i-j||_1} \; , \tag{2.3.32}$$

where $||.||_1$ gives the one-norm (Manhattan metric) of the displacement vector $i-j$. One can show that all diagrams vanish if two nodes are connected to each other

by at least three independent paths. This means, that there are at least three distinctive paths $\rho_{a,j_1}\rho_{j_1 j_2}\ldots\rho_{j_m,b}$ from node (a) to node (b), so that none of the subsegments ρ_{ij} coincide. A trivial example is the diagram in which two internal nodes are connected by at least three independent lines. It will even be possible to eliminate all nontrivial diagrams if the nodes that have a single outgoing and incoming line are eliminated, as shown in Fig. 2.13. For this reason, a gauge in the variational parameters is introduced which sets these nodes to zero

$$Z_{I_1,I_2} = 0 \quad \forall |I_1| = |I_2| = 1 \; . \tag{2.3.33}$$

This is discussed further in section 2.4.

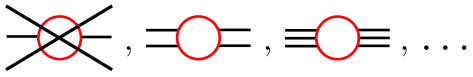


Figure 2.13: The internal nodes with only two lines will be eliminated by a gauge in the variational parameters.

Then, the scaling of the hopping parameters

$$t_{i,j} \sim (\sqrt{2d})^{-||i-j||_1} \tag{2.3.34}$$

shows that all contributions with an internal node or two external nodes that are connected by three or more lines scale at least as $\sim 1/d$. In the limit $d \to \infty$, the only remaining terms are given by

$$\begin{aligned} \langle \hat{T} \rangle_G &= M_{\sigma,\varnothing}(\hat{c}^\dagger_\tau) M_{\varnothing,\sigma'}(\hat{c}_{\tau'}) t^{\tau\tau'}_{ij} \rho^{\sigma\sigma'}_{ij} \\ \langle \hat{U} \rangle_G &= M_{\varnothing}(\hat{U}) \; , \end{aligned} \tag{2.3.35}$$

as shown in Fig.2.14.

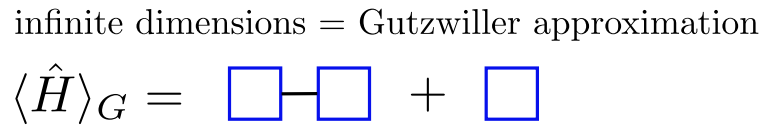


Figure 2.14: Diagrammatic analysis in infinite dimensions.

To lowest order, the coefficients M_{I_1,I_2} coincide with the coefficients Q_{I_1,I_2}

$$\begin{aligned} M_{\varnothing,\sigma'}(\hat{c}_{\tau'}) &= Q_{\varnothing,\sigma'}(\hat{c}_{\tau'}) \; , \\ M_{\varnothing}(\hat{U}) &= Q_{\varnothing}(\hat{U}) \; . \end{aligned} \tag{2.3.36}$$

For $|I_1| + |I_2| = 2$ we have $X_{I_1,I_2} = Z_{I_1,I_2}$. These nodes are used to determine the gauge in the variational parameters where $X_{I_1,I_2} = 0$. The gauge conditions can be computed from the coefficients X_{I_1,I_2} with $I_1 = 1 = I_2$ as it is done in [14]. Note, however, that we must introduce the coefficients Z_{I_1,I_2} to apply the linked cluster theorem. The latter lays the foundations for a further scaling analysis which finally leads to the cancellation of all nontrivial diagrams. Therefore, the analysis is an essential step in the derivation of the exact result (2.3.35) in the limit of infinite dimensions. In the rest of our work we will refer to terms in Eq. (2.3.35) as infinite-d limit.

2.4 Gauge in the variational parameters

In the last section, we derived the exact result for Gutzwiller wave functions on an infinite-dimensional lattice. In this case, the diagrammatic analysis requires that

$$Z_{I_1,I_2} = 0 \quad \forall |I_1| = |I_2| = 1 \; . \tag{2.4.1}$$

For a finite-dimensional system, this gauge will improve our approximation in two ways. It has been shown for the single-band Gutzwiller wave function that this gauge leads to a much faster convergence of the diagrammatic expansion [13]. Furthermore, the gauge condition (2.4.1) ensures the proper behavior of the leading-order term of the expansion. For example, the Gutzwiller correlator reduces to the identity operator for vanishing local interactions. This ensures that the Gutzwiller energy does not violate the Ritz variational principle for vanishing particle interactions.

In principle, the gauge condition imposes a restriction on the variational space because Eq. (2.4.1) implies that not all variational parameter λ_{I_1,I_2} and the single-particle product state $|\Psi_0\rangle$ can be chosen independently. However, we will see that for a special class of Gutzwiller correlators $\hat{P}_G$ and some specific trial states $|\Psi_0\rangle$ the gauge condition does not imply any further variational restrictions. For example, this is the case for our Hubbard model defined in section 3.1. Furthermore, we show that in some cases the full variational freedom might be recovered if it is possible to perform a full variation of the initial trial state $|\Psi_0\rangle$ over all SPPS's. Some final remarks on the rescaling procedure conclude this section.

2.4.1 Rescaling the Gutzwiller correlator

It is illustrative to introduce the following example before we discuss the rescaling procedure for a general Gutzwiller correlator in subsection 2.4.3. Consider a single-band Hubbard model and a Gutzwiller correlator that has only a diagonal single-particle part λ^1, where $(\lambda^1)_{I_1,I_2} = \lambda_{I_1,I_2}$ with $|I_1| = 1 = |I_2|$. As a first step, we can rewrite the Gutzwiller correlator as

$$\hat{P} = \hat{P}_{\text{res}}\hat{\eta} \tag{2.4.2}$$

with

$$\hat{\eta} = \exp(\sum_{i,\sigma} \mu_\sigma \hat{n}_{i,\sigma}) \; . \tag{2.4.3}$$

Then, the rescaled Gutzwiller reads

$$\hat{P}_{\text{res}} = \hat{P} \exp(-\sum_{i,\sigma} \mu_\sigma \hat{n}_{i,\sigma}) \; . \tag{2.4.4}$$

The original wave function is mapped to

$$|\Phi_0\rangle = \prod_{i\sigma} \exp(\mu_\sigma \hat{n}_{i\sigma})|\Psi_0\rangle = \prod_{k\sigma} \exp(\mu_\sigma \hat{n}_{k\sigma})|\Psi_0\rangle \; , \tag{2.4.5}$$

which follows directly from the spatial invariance of the coefficients μ_σ. If $|\Psi_0\rangle$ is an eigenfunction of the total particle number operators $\hat{N}_\sigma = \sum_l \hat{n}_{l\sigma}$, the correlators will merely rescale $|\Psi_0\rangle$ with a scalar factor which drops out after the normalization. Then we can use the two additional parameters to fulfill the gauge constraints. However, this will be the case if and only if $U^k_{\gamma\sigma}$ is already diagonal, or both μ_σ are equal and particle number conservation is guaranteed. This trivial rescaling procedure was already used in [9] and will also be applied in later parts of this work.

In general, $\hat{\eta}$ will change the SPPS $|\Psi_0\rangle$, but we can show that it is mapped to a new SPPS $|\Phi_0\rangle$. To this end, we write down the SPPS explicitly

$$|\Psi_0\rangle = \prod_{k,\gamma_k}^{\text{occ.}} h^\dagger_{k\gamma_k}|\text{vac}\rangle = \prod_k |\psi_k\rangle \,, \tag{2.4.6}$$

$$h^\dagger_{k\gamma} = \bar{U}^{ki}_{\gamma\sigma} c^\dagger_{i\sigma} \,, \quad c^\dagger_{i\sigma} = U^{ki}_{\gamma\sigma} h^\dagger_{k\gamma} \,, \tag{2.4.7}$$

where the bar over the transformation tensor U denotes the complex conjugation. For a translational invariant system this tensor can be decomposed into

$$U^{ki}_{\gamma\sigma} = U^k_{\gamma\sigma} V^k_i \,, \tag{2.4.8}$$

where $V^k_i \sim \exp(\mathrm{i}ki)$ gives the usual Fourier transformation and $U^k_{\gamma\sigma}$ gives the transformation matrix obtained in the diagonalization of the Hamilton operator for fixed momentum k. The action of $\hat{\eta}$ on $|\Psi_0\rangle$ will lead to the transformation

$$\tilde{U}^{ki}_{\gamma\sigma} = U^{ki}_{\gamma\sigma} e^{\mu_\sigma} \,, \quad \tilde{U}^k_{\gamma\sigma} = U^k_{\gamma\sigma} e^{\mu_\sigma} \,, \tag{2.4.9}$$

where the second term simplifies due to translation symmetry. Then, the new wave function can be written as

$$|\Phi_0\rangle = \prod_{k\gamma_k}^{\text{occ.}} \bar{U}^{ki}_{\gamma\sigma} e^{\mu_\sigma} \hat{c}^\dagger_\sigma |\text{vac}\rangle = \prod_{k\gamma_k}^{\text{occ.}} \tilde{h}^\dagger_{k\gamma_k}|\text{vac}\rangle = \prod_k |\phi_k\rangle \,, \tag{2.4.10}$$

with

$$\tilde{h}^\dagger_{k\gamma} = \bar{U}^k_{\gamma\sigma} e^{\mu_\sigma} \hat{c}^\dagger_{k\sigma} = \bar{U}^k_{\gamma\sigma} e^{\mu_\sigma} U^k_{\gamma'\sigma} \hat{h}^\dagger_{k\gamma'} \,. \tag{2.4.11}$$

The transformation between the ladder operators can be written as

$$\tilde{h}^\dagger_{k\gamma} = \bar{U}^k_{\gamma\sigma} e^{\mu_\sigma} c^\dagger_{k\sigma} = \bar{\tilde{U}}^k_{\gamma\sigma} c^\dagger_{k\sigma} \,, \tag{2.4.12}$$

$$c^\dagger_{i\sigma} = e^{-\mu_\sigma} U^{ki}_{\gamma\sigma} \tilde{h}^\dagger_{k\gamma} = \tilde{U}^k_{\gamma\sigma} \tilde{h}^\dagger_{k\gamma} \,. \tag{2.4.13}$$

Wick's theorem applies only for single-particle product states. Even though we can still write $|\Phi\rangle$ as a product of operators which act on the vacuum state, these operators do not fulfill the canonical commutation relations which become

$$\{\tilde{h}^\dagger_{k\gamma}, \tilde{h}_{k'\gamma'}\} = \sum_{kk',\sigma\sigma'} \bar{\tilde{U}}^k_{\gamma\sigma} \tilde{U}^{k'}_{\gamma'\sigma'} \{c^\dagger_{k\sigma}, c_{k\sigma'}\} \tag{2.4.14}$$

$$= \sum_\sigma e^{2\mu_\sigma} \bar{U}^k_{\gamma\sigma} U^k_{\gamma'\sigma} \,. \tag{2.4.15}$$

In the next subsection, we will transform the operators $\tilde{h}^\dagger_{k\gamma}$ in such a way that the anticommutation relations are restored.

2.4.2 Reorthogonalisation

For a fixed momentum k, we can perform a QR-decomposition of the matrix $\tilde{U}^k_{\gamma\sigma}$ in Eq. (2.4.12),

$$\tilde{U} = U\eta^1 = LQ^T = (QR)^T \,, \tag{2.4.16}$$

$$\tilde{U}^\dagger \tilde{U} = (\eta^1)^\dagger \eta^1 \,, \tag{2.4.17}$$

$$\tilde{U}\tilde{U}^\dagger = LL^\dagger \,, \tag{2.4.18}$$

where Q is unitary and L and R is a lower and upper triangular matrix, respectively. For clarity of notation, we dropped the momentum index. The details of the QR-decomposition can be found in any textbook on numerical analysis, e.g. [17]. The matrix Q can be used to define new Fermi operators

$$\hat{f}_{k\tau} = \sum_\sigma Q^k_{\tau,\sigma} \hat{h}_{k\sigma} \,, \tag{2.4.19}$$

such that

$$|\phi_k\rangle = \prod_{\gamma_k}^{\text{occ.}} L^k_{\gamma\tau} \hat{f}_{k\tau} |\text{vac}\rangle \,. \tag{2.4.20}$$

Let us assume that we reordered the index γ in such that the first n occupied levels obeys $\gamma_1 < \ldots < \gamma_n$. Then, we find

$$|\phi_k\rangle = \prod_{\tau_k}^{\text{occ.}} L^k_{\tau\tau} \hat{f}_{k\tau} |\text{vac}\rangle \,, \tag{2.4.21}$$

where the lower triangular form of L and the Fermi character of the operators guarantee that only the diagonal entries of L contribute. The product of the diagonal entries will cancel out after the normalization. The momentum representation of the density matrix $\rho^k_{\sigma\sigma'}$ in Eq. (2.1.21) can now be evaluated by an exchange of the matrices U and Q

$$\rho^k_{\sigma\sigma'} = \sum_{\gamma_k}^{\text{occ.}} \bar{Q}^k_{\gamma\sigma} Q^k_{\gamma\sigma'} \,. \tag{2.4.22}$$

This procedure can only be applied if the ordering of the band index γ is chosen in such a way that the occupied states have the lowest index for all k. This can also be understood if we think of Q as the matrix which contains the orthogonalized vectors obtained by a usual Gram-Schmidt decomposition. Then, the choice of the orthogonal system depends on the sequence of the orthogonalization procedure. Typically, the band index is ordered in the same way as the energies so that this constraint is fulfilled for a Fermi state. Therefore, the reorthogonalization of $|\Psi_0\rangle$ leads to a new SPPS

$$\hat{\eta}|\Psi_0\rangle = \alpha|\Phi_0\rangle \tag{2.4.23}$$

with some constant α in front of the new wave function that will drop out when the denominator of the expectation value is taken into account.

In the simplest cases, we can set $\eta^1 \sim$ Id (or U and η^1 are both diagonal). Then, we can set $L = \eta^1$ and $Q = U$ which shows that there are no nondiagonal entries in L, so that the density matrix will not change. In this section, we focused on a translational invariant system. Nevertheless, the reorthogonalization could also be performed for a system with a site dependent $\eta^1_{l,\sigma\sigma'}$. Then, the matrices in Eq. (2.4.16) would have to be defined for the whole system. Note, that the reorthogonalization procedure described above is just an adaption of the one given in [14].

In general, the reorthogonalization will change the gauge conditions which depend on the local entries of the new density matrix ρ_{Φ_0}. Therefore, we will analyze the following example in order to examine the structure of ρ_{Φ_0}. We consider the action of the Gutzwiller correlator on the ground state of the following kinetic energy operator of a translational invariant system,

$$\hat{T}^k = \begin{pmatrix} \epsilon^k_{11} & \epsilon^k_{12} \\ \epsilon^k_{12} & \epsilon^k_{22} \end{pmatrix} , \tag{2.4.24}$$

with $\epsilon_1 = t_{11}(\cos(kx) + \cos(ky))$, $\epsilon_2 = t_{22}(\cos(kx) + \cos(ky))$ and $\epsilon_{12} = t_{12}(\cos(kx) - \cos(ky))$ where the $t_{\sigma\sigma'}$ denote the different nearest-neighbor hopping parameter. The index $\sigma = 1, 2$ can be regarded as the spin index or the band index of a system without spin flips. Furthermore, a reflection through the x and y axis shall lead to the transformation $\epsilon_{12} \to -\epsilon_{12}$ while the inversion $k \to -k$ leaves $\hat{T}^k$ invariant. The transformation matrix can be written as

$$U^k = \begin{pmatrix} \cos(\phi_k) & \sin(\phi_k) \\ -\sin(\phi_k) & \cos(\phi_k) \end{pmatrix} , \phi_k = \frac{2\epsilon^k_{12}}{\epsilon^k_1 - \epsilon^k_2} , \tag{2.4.25}$$

with $\sin(\phi_k) = s_k$ and $\cos(\phi_k) = c_k$. Then,

$$\rho^k = \begin{pmatrix} c_k^2 & s_k c_k \\ s_k c_k & s_k^2 \end{pmatrix} \theta^k_1 + \begin{pmatrix} s_k^2 & -s_k c_k \\ -s_k c_k & c_k^2 \end{pmatrix} \theta^k_2 . \tag{2.4.26}$$

In order to calculate the local density matrix $\rho_{\sigma\sigma'}$ we sum over all four symmetry points $[k] = (\pm k_x, \pm k_y)$. At this point, we will not discuss the detail of the calculations and simply state the result

$$\sum_{k\in[k]} \rho^k = \rho^k = \begin{pmatrix} 2(c_k^2\theta^k_1 + s_k^2\theta^k_2) & 0 \\ 0 & 2(s_k^2\theta^k_1 + c_k^2\theta^k_2) \end{pmatrix} . \tag{2.4.27}$$

After the reorthogonalization, the density matrix transforms to

$$\begin{aligned} \tilde{\rho}^k = &\begin{pmatrix} (c_k b_1)^2 & s_k c_k b_1 b_2 \\ s_k c_k b_1 b_2 & (s_k b_2)^2 \end{pmatrix} \frac{\theta^k_1(1-\theta^k_2)}{(c_k b_1)^2 + (s_k b_2)^2} \\ &+ \begin{pmatrix} (s_k b_2)^2 & -s_k c_k b_1 b_2 \\ -s_k c_k b_1 b_2 & (c_k b_1)^2 \end{pmatrix} \frac{\theta^k_2(1-\theta^k_1)}{(s_k b_2)^2 + (c_k b_1)^2} + \theta^k_1\theta^k_2 \text{ Id} \end{aligned} \tag{2.4.28}$$

with $b_i = \exp(\mu_i)$, and

$$\sum_{k\in[k]} \tilde{\rho}^k = \begin{pmatrix} (c_k b_1)^2 & 0 \\ 0 & (s_k b_2)^2 \end{pmatrix} \frac{2\theta_1^k(1-\theta_2^k)}{(c_k b_1)^2 + (s_k b_2)^2} + \begin{pmatrix} (s_k b_2)^2 & 0 \\ 0 & (c_k b_1)^2 \end{pmatrix} \frac{2\theta_2^k(1-\theta_1^k)}{(s_k b_2)^2 + (c_k b_1)^2} + 4\theta_1^k\theta_2^k \text{ Id} . \tag{2.4.29}$$

The diagonality of $\tilde{\rho}^k$ is still preserved by the symmetries of $\hat{T}$ and the Fermi surface is not changed by the reorthogonalization procedure. However, we have to distinguish if the first band, the second band, or both bands are occupied. This shows that long-ranged correlations are build into $\rho_{\Phi_0} = \tilde{\rho}_{ij}^{\sigma\sigma'}$ to incorporate the nontrivial k dependency. Additionally, it is easy to see that we cannot find a simple expression for the new densities $\tilde{n}_1$ and $\tilde{n}_2$. The coefficients Z_{I_1,I_2} are now complicated functions of the rescaling parameters μ_1 and μ_2 and a self-consistent solution of μ_1 and μ_2 (if existent) will only be possible with numerical methods. In order to avoid an exact evaluation of the mapping $|\Psi_0\rangle \to |\Phi_0\rangle$, we can treat $|\Phi_0\rangle$ as variational quantity. In this way, we can recover the full variational freedom as long as a self-consistent solution formally exists. This will be shown in the next section for the general multi-band case.

2.4.3 Generalized rescaling procedure

In the following section, we will focus on a Gutzwiller operator $\hat{P}$ that does not mix states with a different number of particles. In order to fulfill the gauge condition, Eq. (2.4.1), we need to rescale the variational parameters $(\lambda^1)_{I_1,I_2} = \lambda_{I_1,I_2}$ with $|I_1| = 1 = |I_2|$. This is possible if we can decompose the matrix λ^1 in the following way

$$\lambda^1 = \lambda^1_{\text{res}}\, \eta^1 , \tag{2.4.30}$$

where λ^1_{res} gives the desired values of the variational parameters. The matrix η^1 needs to be invertible and Hermitian. The action of η^1 on a single-particle state $\hat{c}^\dagger_\sigma|\text{vac}\rangle$ will result in a transformation of the single particle basis.

$$\sum_{\substack{I_1,I_2 \\ |I_1|=1=|I_2|}} (\eta^1)_{I_1,I_2} |I_1\rangle\langle I_2|\, \hat{c}^\dagger_\sigma|\text{vac}\rangle = \eta_{\sigma'\sigma}\hat{c}^\dagger_{\sigma'}|\text{vac}\rangle . \tag{2.4.31}$$

Now we need to expand η^1 to the many-particle sector in such a way that the same single-particle transformation is applied. Therefore, we must construct the remaining entries of the full matrix η with the help of the embedding ι defined in Eq. (2.1.25),

$$\eta_{I_1,I_2} = \det[\tilde{\eta}_{I_1,I_2}] , \tag{2.4.32}$$

where $\tilde{\eta}$ gives the submatrix consisting of the elements $\eta^1_{\sigma_1,\sigma_2}$ $\sigma_1 \in I_1$ and $\sigma_2 \in I_2$ with standard ordering, and $\eta_\varnothing = 1$. The inverse of this matrix is given by the

corresponding many-particle extension of the inverse of η^1. Now, we can split up the Gutzwiller correlator into two parts

$$\hat{P} = \hat{P}_{\text{res}}\, \hat{\eta}\ . \tag{2.4.33}$$

$\hat{P}_{\text{res}}$ has the desired parameters in the single-particle sector. The new coefficients are given by

$$\lambda_{\text{res}} = \lambda\, \eta^{-1}\ . \tag{2.4.34}$$

We are still free to perform a variation of all parameters in λ with $|I_1| = |I_2| > 1$ so that the coefficients in λ_{res} can still be considered as independent parameters with λ^1_{res} fixed to the desired value. The action of $\hat{\eta}$ on the single-particle product state $|\Psi_0\rangle$ will lead to a transformation of the single-particle basis so that we obtain the new SPPS

$$\hat{\eta}|\Psi_0\rangle = \alpha|\Phi_0\rangle\ . \tag{2.4.35}$$

The constant α in front of the new wave function will drop out when the denominator of the expectation value is taken into account.

We may hope that a matrix η can always be found so that the rescaled variational parameters fulfill the gauge constraints. In general, the new state $|\Phi_0\rangle$ will lead to a different set of gauge conditions which depend on the local entries of the new density matrix ρ_{Φ_0}, as it was shown in the last subsection. Therefore, we need to find a self-consistent solution of η. Let us assume that we can always find such a self-consistent solution. Then, the mapping $(\lambda, |\Psi_0\rangle) \to (\lambda_{\text{res}}, |\Phi_0\rangle)$ becomes surjective. If we find the optimal solution of the right-hand-side of

$$\hat{P}(\lambda)|\Psi\rangle = \hat{P}(\lambda_{\text{res}})|\Phi_0\rangle\ , \tag{2.4.36}$$

we have solved the full minimization problem. The foregoing example shows that it is very difficult to construct an explicit solution of the reothogonalization procedure. An optimization algorithm that covers the full variational space of $|\Phi_0\rangle$, however, automatically includes the reorthogonalization procedure and the changes in ρ_Φ. Therefore, we can compensate the loss in the variational freedom of the Gutzwiller correlator by a variation of the SPPS.

A similar argumentation was used for a diagonal Gutzwiller correlator in the single-band case in [10]. It has been argued that a variation over the parameters and wave functions in

$$\prod_l \exp(1 + \mu_d \hat{n}_{\uparrow,l}\hat{n}_{\downarrow,l})|\Psi_0\rangle \tag{2.4.37}$$

and

$$\prod_l \left(\lambda_0|\varnothing\rangle\langle\varnothing| + \lambda_\uparrow|\uparrow\rangle\langle\uparrow| + \lambda_\downarrow|\downarrow\rangle\langle\downarrow| + \lambda_d|\uparrow\downarrow\rangle\langle\uparrow\downarrow|\right)_l |\Psi_0\rangle \tag{2.4.38}$$

cover the same variational space. In [13], this relation has been imprecisely used by the authors to argue that the restrictions of the gauge constraints in Eq. (2.4.1) can

be lifted. This, however, does not guarantee that we can find a self-consistent solution for η. As it was shown in subsection 2.4.1, this assumption does, however, hold for a paramagnetic wave function $|\Psi_0\rangle$ with densities $n^0_\uparrow = n^0_\downarrow$ which is considered by the authors in [13]. In this case a self consistent solution can be trivially found because $\lambda^1 \sim$ Id and $\rho_{\Psi_0} = \rho_{\Phi_0}$.

2.4.4 Remarks on the choice of the variational parameters

The original Gutzwiller correlator should be Hermitian $\lambda = \lambda^\dagger$ for most physical setups. Intuitively, one might assume that the rescaled Gutzwiller correlator, which is the one which will be used in the optimization procedure, is also Hermitian $\lambda_{\text{res}} = \lambda^\dagger_{\text{res}}$. The matrix η is defined as a single particle transformation and therefore always Hermitian. This requires that

$$[\lambda_{\text{res}}, \eta] = 0 \; . \tag{2.4.39}$$

Thus, we have to demand that λ, λ_{res} and η are simultaneously diagonalizable. Since the Gutzwiller operator does not mix states with a different number of particles, η is uniquely defined by its single-particle sector which can be digonalized by the transformation matrix V. The unitary matrix V will define the matrix $W = \iota(V)$ via the embedding defined in Eq. (2.1.25). Then, the matrix W will diagonalize the multi-particle components of η, λ and λ_{res}. However, it is clear that we will not be able to diagonalize the Gutzwiller correlator with a single-particle transformation if it includes more complicated many-particle correlations. In this case we would have to drop the restriction that λ_{res} is Hermitian. For example, we will have to use a non-Hermitian set of coefficients for λ_{res} if the single band Gutzwiller correlator is nondiagonal and $\hat{\eta} = \exp(\mu_1 \hat{n}_1 + \mu_2 \hat{n}_1)$, with $\mu_1 \neq \mu_2$. In the next section we will define a two-band Hubbard model with a nondiagonal Gutzwiller correlator. In this case, we would have to use a non-Hermitian set of variational parameters if we allowed different orbital densities.

Another problem can occur if we start with a diagonal local density matrix ρ and diagonal coefficients η, λ and λ_{res}. The new single particle correlation matrix $\tilde{\rho}$ has to remain diagonal to avoid the emergence of further gauge constraints. However, the diagonality of the density matrix is often guaranteed by the symmetries of the system, and will most likely not be affected by the rescaling procedure.

Chapter 3

Two-orbital Hubbard model on a square lattice

In this chapter, we evaluate the Gutzwiller wave function for a two-orbital Hubbard model on a square lattice. The diagrammatic expansion is carried out to second order. Details of the computational generation of the diagrams and the optimization routines of the Gutzwiller variational parameters can be found in appendix A.2.

In the first section, we derive the two-band Hamilton operator for p_x-p_y-type orbitals on a square lattice. We derive the lattice symmetries and incorporate them into our Gutzwiller wave function. In the second section, we analyze the ferromagnetic phase transition in this model. We show that a spin polarization of the Gutzwiller wave function requires much larger electron interactions than we would expect from Hartree-Fock theory. Furthermore, the regions of finite but nonsaturated ferromagnetism cover a much larger parameter space in Gutzwiller theory. In the third section, we show that the correlations lead to a deformation of the Fermi surface of the underlying single-particle product state. The deformations are clearly visible and can even lead to a change of the Fermi-surface topology. Furthermore, we compare our method to previous results of a higher-order study of the single-band Hubbard model. Moreover, we show that the Fermi surface deformations can break the lattice symmetry in this model (Pomeranchuk instability).

3.1 Model specifications

In this section, we derive the two-band Hamilton operator for p_x-p_y-type orbitals on a square lattice. As an application, this model describes fermions trapped in a two-dimensional optical lattice. Some work on the realization of p_x-p_y orbitals in optical lattices can be found in [18–20]. The ground-state properties of this model are still subject of current research. For example, some analytical work on a p_x-p_y Hubbard model on square and cubic lattices can be found in [21]. Some approaches that are related to our Gutzwiller variational wave function have been applied to similar models. For example, a study of the single-band Hubbard model within an extended

Gutzwiller approach can be found in [22]. The same model is investigated in [23, 24] and [25], where a variational Monte Carlo approach is used that employs a nonlocal Gutzwiller-Jastrow correlator. A study of a two-dimensional bilayer Hubbard model with the same method can be found in [26].

In this section we first introduce our Hubbard model with electron transfers between nearest and next-nearest neighbors, and derive the symmetry relations and the resulting structure of the density matrix $\rho_{ij}^{\sigma,\sigma'}$. The local interaction energy can be expressed by two independent model parameters. Next, we implement the symmetries in the Gutzwiller correlator, and consider the corresponding effects on the internal and external nodes of our diagrammatic analysis. Finally, we comment on the applicability of our model for a set of d_{xz}-d_{yz} orbitals.

3.1.1 Operator for the kinetic energy

The hopping amplitudes $t_{ij}^{\sigma\sigma'}$ of the kinetic energy operator $\hat{T} = \sum_{ij} t_{ij}^{\sigma\sigma'} \hat{c}_{i\sigma}^{\dagger} \hat{c}_{j\sigma'}$ are considered to be free model parameters in our work. However, they are restricted by the rotational symmetries of p_x-p_y orbitals. For simplicity, we drop the spin index for the moment because we do not allow any spin-flip terms in our Hamilton operator.

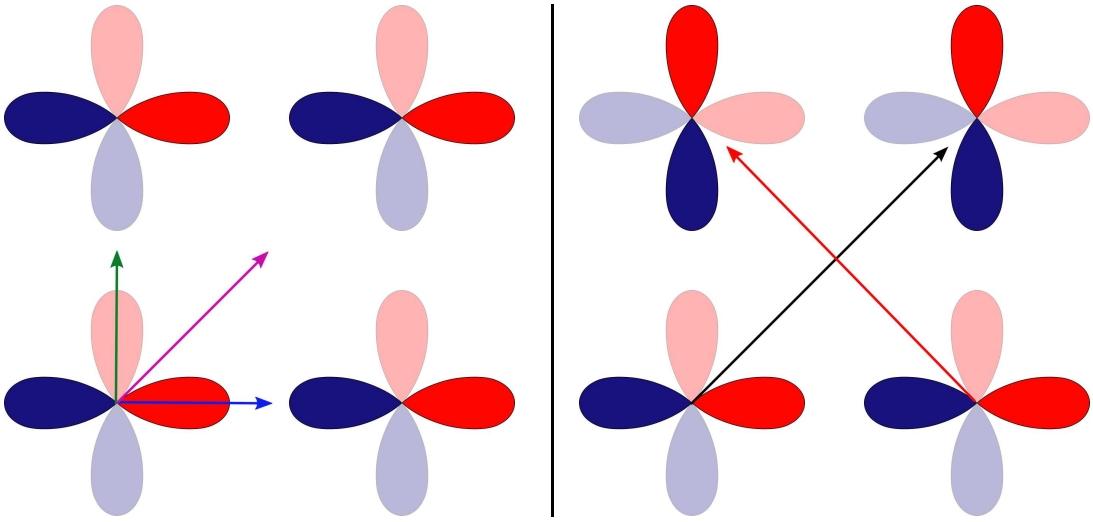


Figure 3.1: Hopping amplitudes to nearest and next-nearest neighbors. The p_x and p_y orbitals are aligned horizontally and vertically, respectively. A red color indicates a positive sign of the wave function while a blue color indicates a negative sign. Left: Hopping amplitudes for the transitions between the p_x orbitals. A rotation by $\pi/2$ gives the hopping amplitudes for the transitions between the p_y orbitals. Right: Hopping amplitudes for the transitions between the p_x and p_y orbitals. The sign of the hopping amplitude changes after a rotation of $\pi/2$. Transitions to nearest neighbors are forbidden by the symmetry of the orbitals.

Fig. 3.1 shows the hopping processes to nearest and next-nearest neighbors. The amplitude t_x^{11} for the transition from the p_x ($\tau = 1$) orbitals on site i to the site $i \pm \mathrm{d}x$ equals the amplitude t_y^{22} for the transition from the p_y orbitals ($\tau = 2$) on

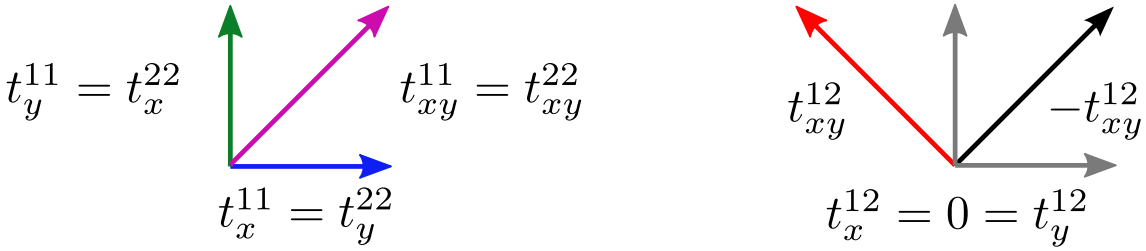


Figure 3.2: Hopping Amplitudes for the p_x-p_x model. Left: The transition processes within an orbit are described by three model parameters. Right: The transition between p_x and p_y is described by one free parameter.

site i to the site $i \pm \mathrm{d}y$. The same holds for the amplitudes t_{xy}^{11} and t_{xy}^{22} for the transitions between the p_x orbitals on i and $i \pm \mathrm{d}y$ and the p_y orbitals on sites i and $i \pm \mathrm{d}x$ respectively. The amplitudes for the hopping processes from i to $i \pm \mathrm{d}x \pm \mathrm{d}y$ between the p_x orbitals is the same as between the p_y orbitals, see Fig. 3.1.

The symmetry of the orbitals does not allow any p_x-p_y transition to nearest neighbors. For transitions between next-nearest neighbors, the sign of the amplitudes t_{xy}^{12} will change after a rotation of $\pi/2$ so that $t_{xy}^{12} = -t_{yx}^{12}$. Fig. 3.2 shows all four remaining free parameters. The xy symmetry of the inter-orbital hopping processes leads to a diagonal local density matrix $\rho_{\sigma,\sigma'} = n_\sigma^0 \delta_{\sigma,\sigma'}$, as we shown in the next subsection. Furthermore, the rotational symmetry of the hopping processes within the same orbitals guarantees that all diagonal entries of the local density matrix are the same, $\rho_{\sigma,\sigma'} = n_0$ Id. This property allows a simple gauge in the variational parameters of the Gutzwiller wave function, as discussed in section 2.4.

The Hamilton operator displays particle-hole symmetry as discussed further in appendix A.3.4. A system with density n_0 corresponds to a system with $\tilde{n}_0 = 1 - n_0$, if the sign of the next-nearest neighbor t_{xy}^{11} hopping is reversed. In general, the numerical calculations become more cumbersome for larger densities so that this mapping allows us to restrict ourselves to densities below half filling.

3.1.2 Density matrix

In momentum space, the matrix representation of the kinetic energy operator reads

$$\epsilon(k) = \begin{pmatrix} \epsilon_{11}^k & \epsilon_{12}^k \\ \epsilon_{12}^k & \epsilon_{22}^k \end{pmatrix} , \tag{3.1.1}$$

with momentum k and

$$\epsilon_1^k = t_x^{11}(\cos(k_x) + \cos(k_y)) + t_{xy}^{11}(\cos(k_x + k_y) + \cos(k_x - k_y)) , \tag{3.1.2}$$
$$\epsilon_2^k = t_x^{22}(\cos(k_x) + \cos(k_y)) + t_{xy}^{11}(\cos(k_x + k_y) + \cos(k_x - k_y)) , \tag{3.1.3}$$
$$\epsilon_{12}^k = t_{xy}^{12}(\cos(k_x + k_y) - \cos(k_x - k_y)) . \tag{3.1.4}$$

In section 3.3 we use an effective Hamilton operator $\hat{T}_{\text{eff}}$ for the computation of the SPPS $|\Psi_0\rangle$. Even so, $\epsilon(k)$ will have the same symmetries. The ground state can be

constructed by a diagonalization, $\epsilon_{\tau,\tau'}(k) = \bar{U}_{\gamma\tau}\epsilon_\gamma U^k_{\gamma\tau'}$ so that

$$|\Psi_0\rangle = \prod_{k,\gamma_k}^{\text{occ.}} h^\dagger_{k\gamma_k}|\text{vac}\rangle \, , h^\dagger_{k\gamma} = \bar{U}^k_{\gamma\tau} c^\dagger_{k\tau} \, . \tag{3.1.5}$$

The transformation matrix can be written as

$$U^k = \begin{pmatrix} \cos(\phi_k) & \sin(\phi_k) \\ -\sin(\phi_k) & \cos(\phi_k) \end{pmatrix} \, , \, \tan(2\phi_k) = \frac{2\epsilon^k_{12}}{\epsilon^k_{11} - \epsilon^k_{22}} \, . \tag{3.1.6}$$

The eigenenergies are given by

$$\epsilon_1(k) = \epsilon^k_{11}\cos(\phi_k)^2 + \epsilon^k_{22}\sin(\phi_k)^2 + \epsilon^k_{12}\sin(2\phi_k) \, , \tag{3.1.7}$$

$$\epsilon_2(k) = \epsilon^k_{22}\cos(\phi_k)^2 + \epsilon^k_{11}\sin(\phi_k)^2 - \epsilon^k_{12}\sin(2\phi_k) \, . \tag{3.1.8}$$

The Fourier-transformed density matrix becomes

$$\rho^k = \begin{pmatrix} c_k^2 & s_k c_k \\ s_k c_k & s_k^2 \end{pmatrix} \theta^k_1 + \begin{pmatrix} s_k^2 & -s_k c_k \\ -s_k c_k & c_k^2 \end{pmatrix} \theta^k_2 \, , \tag{3.1.9}$$

where $\theta^k_i = \theta(E_\text{F} - \epsilon_i(k))$ is the usual Heaviside step function, $\sin(\phi_k) = s_k$ and $\cos(k) = c_k$. The symmetries of the hopping amplitudes are passed over to the density matrix. For example a rotation of k by $\pi/2$ maps $s_k \to -s_k$. Therefore, the symmetry relations of $t^{\sigma\sigma'}_{ij}$ also apply to $\rho^{\sigma\sigma'}_{ij}$.

Now we can prove the propositions about the local density matrix $\rho^{\sigma\sigma'}_{ii}$ made in the previous subsection. The nondiagonal part ϵ_{12} changes its sign after a reflection $k_x \to -k_y$ while ϵ^k_{11} and ϵ^k_{22} remain unchanged. The eigenenergies $\epsilon_1(k)$ and $\epsilon_2(k)$ are not affected. The sum over the points (k_x, k_y) and $(-k_x, k_y)$ yields a diagonal density matrix

$$\rho^{(k_x,k_y)} + \rho^{(-k_x,k_y)} = \begin{pmatrix} c_k^2\theta^k_1 + s_k^2\theta^k_2 & 0 \\ 0 & s_k^2\theta^k_1 + c_k^2\theta^k_2 \end{pmatrix} \, . \tag{3.1.10}$$

A reflection in the x-axis interchanges the role of ϵ^k_{11} and ϵ^k_{22} but ϵ^k_{12} stays unaltered this time. ϕ_k changes its sign so that ϵ_1 and ϵ_2 are interchanged. The eigenvector for the larger (smaller) eigenvalue of $\epsilon(k)$ now corresponds to the smaller (larger) eigenvalue. Therefore, we can obtain the new density matrix by interchanging θ^k_1 and θ^k_2 and adjusting the sign of s_k. Then,

$$\rho^{(k_x,k_y)} + \rho^{(k_y,k_x)} = \begin{pmatrix} 1 & 2s_k c_k \\ 2s_k c_k & 1 \end{pmatrix} \theta^k_1 + \begin{pmatrix} 1 & -2s_k c_k \\ -2s_k c_k & 1 \end{pmatrix} \theta^k_2 \, , \tag{3.1.11}$$

which shows that the diagonal entries of the density matrix are the same. The calculation of $\rho^{\sigma\sigma'}_{ii}$ invokes the summation over all possible symmetry points, namely (k_x, k_y),(k_y, k_x), $(-k_x, k_y)$,$(-k_y, k_x)$, $(k_x, -k_y)$,$(k_y, -k_x)$, $(-k_x, -k_y)$, and $(-k_y, -k_x)$, generated by the D_4 lattice symmetry. The sum over all points yields

$$\rho^{[k]} = 8 \, \text{Id}(\theta^k_1 + \theta^k_2) \, , \tag{3.1.12}$$

where the weight must be adjusted if k lies on a symmetry axis.

3.1.3 Local interaction energy

The calculation of the on-site Coulomb interaction

$$\hat{U}_{\text{int}} = \sum_{\sigma_1,\sigma_2,\sigma_3,\sigma_4} U_{\sigma_1,\sigma_2,\sigma_3,\sigma_4} \hat{c}^\dagger_{\sigma_1} \hat{c}^\dagger_{\sigma_2} \hat{c}_{\sigma_3} \hat{c}_{\sigma_4} \tag{3.1.13}$$

requires the evaluation of two-particle expectation values of the Coulomb energy

$$I^{s_2 s_3}_{s_1 s_4} = \langle \phi_{s_1} \phi_{s_2} | \hat{V}_{\text{Coul.}} | \phi_{s_3} \phi_{s_4} \rangle \, . \tag{3.1.14}$$

Here, the atomic orbitals ϕ_s with multi-index $s = (n, l, m)$ are given by the wave functions

$$\phi_{n,l,m}(r, \theta, \varphi) = R_{nl}(r) Y_{lm}(\theta, \varphi) \, , \tag{3.1.15}$$

where n is the principal quantum number and l the absolute value of the angular momentum and m its z-component. The wave function can be split up into a radial part $R_{nl}(r)$ and an angular dependent part $Y_{lm}(\theta, \varphi)$ that are defined in terms of Laguerre and Legendre polynomials, respectively. The following derivations can be found in many text books (e.g. [27]) but are repeated here to illustrate the concept. The expectation values in (3.1.14) can be expressed by the integrals

$$\begin{aligned} I^{s_2 s_3}_{s_1 s_4} &= \int\int \phi^*_{s_1}(r_a) \phi^*_{s_2}(r_b) \frac{1}{r_{ab}} \phi_{s_3}(r_b) \phi_{s_4}(r_a) \, \mathrm{d}V_a \, \mathrm{d}V_b \\ &= \int\int Y^*_{s_1}(\theta_a, \varphi_a) Y^*_{s_2}(\theta_b, \varphi_b) Y_{s_3}(\theta_b, \varphi_b) Y_{s_4}(\theta_a, \varphi_a) \frac{R^2(r_a) R^2(r_b)}{r_{ab}} \, \mathrm{d}V_a \, \mathrm{d}V_b \, , \end{aligned} \tag{3.1.16}$$

where the electric charge and dielectric constant have been set to unity. The distance vector r_{ab} can be expanded in spherical harmonics so that the integrals can be simplified to

$$I^{s_2 s_3}_{s_1 s_4} = \sum_{l=0}^{\infty} SC(l) L_l(l_{s_1}, m_{s_1}; l_{s2}, m_{s_2}; l_{s_3}, m_{s_3}; l_{s_4}, m_{s_4}) \tag{3.1.17}$$

with some coefficients L and the Slater–Condon integrals

$$SC(l) = 4\pi \int \mathrm{d}r_a \int \mathrm{d}r_b \frac{\min(r_a, r_b)^l}{\max(r_a, r_b)^{l+1}} R^2(r_a) R^2(r_b) \, . \tag{3.1.18}$$

The coefficients L can be evaluated exactly. They are given in appendix A.3. In our case, only the Slater–Condon integrals for $l = 0, 2, 4$ are required. In principle, all entries in $I_{s_1 s_4; s_2 s_3}$ are linear combinations of the three coefficients $SC(0)$, $SC(2)$ and $SC(4)$ (or Racah parameters A, B, C) [27]. The explicit calculation of the coefficients L shows that we need only two independent Hubbard parameters U and J. The matrix representation of the two-particle sector of $\hat{U}_{\text{int}}$ is given by

$$U_{\text{int}} = U \, \text{Id} + \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & J \\ 0 & -3J & 0 & 0 & 0 & 0 \\ 0 & 0 & -2J & -J & 0 & 0 \\ 0 & 0 & -J & -2J & 0 & 0 \\ 0 & 0 & 0 & 0 & -3J & 0 \\ J & 0 & 0 & 0 & 0 & 0 \end{pmatrix} , \tag{3.1.19}$$

where the standard ordering $\{|\uparrow\downarrow,0\rangle, |\uparrow,\uparrow\rangle, |\downarrow,\uparrow\rangle, |\uparrow,\downarrow\rangle, |\downarrow,\downarrow\rangle, |0,\uparrow\downarrow\rangle\}$ of the two-particle states has been used. Spin conservation and the conservation of the z-component of the angular momentum, $m_1 + m_2 = m_3 + m_4$, is guaranteed. The latter constraint excludes any terms where only one electron switches from a p_x to a p_y orbital.

The two-particle sector provides all necessary coefficients $U_{\sigma_1,\sigma_2,\sigma_3,\sigma_4}$. The energy is zero if an atom is empty or singly occupied. The local potential energy operator is diagonal in the three-particle sector because the nondiagonal terms in U_{int} involve operators of all four local configurations. Every triple-occupied state is annihilated by these terms. The energy for a triple or quadruple occupancy are $3U - 5J$ and $2(3U - 5J)$, respectively.

A diagonalization of the two-particle sector of the local potential energy operator gives the following eigenenergies and eigenvectors

$$\begin{aligned}
&|\Gamma_1\rangle = (|\uparrow\downarrow,0\rangle + |0,\uparrow\downarrow\rangle)2^{-1/2} && U_{\Gamma_1} = U + J && S = 0 && S_z = 0 \\
&|\Gamma_2\rangle = |\uparrow,\uparrow\rangle && U_{\Gamma_2} = U - 3J && S = 1 && S_z = 1 \\
&|\Gamma_3\rangle = (|\uparrow,\downarrow\rangle + |\downarrow,\uparrow\rangle)2^{-1/2} && U_{\Gamma_3} = U - 3J && S = 1 && S_z = 0 \\
&|\Gamma_4\rangle = (|\uparrow,\downarrow\rangle - |\downarrow,\uparrow\rangle)2^{-1/2} && U_{\Gamma_4} = U - J && S = 0 && S_z = 0 \\
&|\Gamma_5\rangle = |\downarrow,\downarrow\rangle && U_{\Gamma_5} = U - 3J && S = 1 && S_z = -1 \\
&|\Gamma_6\rangle = (|\uparrow\downarrow,0\rangle - |0,\uparrow\downarrow\rangle)2^{-1/2} && U_{\Gamma_6} = U - J && S = 0 && S_z = 0\,.
\end{aligned} \tag{3.1.20}$$

The eigenvectors form a triplet of degenerate states with energy $U - 3J$ and a doublet with energies $U - J$ and $U + J$, respectively. The states with the highest spin multiplicity are lowest in energy, in agreement with Hund's rule. However, these local symmetries are not conserved by the kinetic energy operator $\hat{T}$. This can be be seen by an examination of the ground state of a system with four sites and periodic boundary conditions for example. Further remarks are given in appendix A.3. The next subsection discusses further implications of these symmetries, and how to incorporate them in the Gutzwiller wave function.

3.1.4 Symmetries in the Gutzwiller wave function

In previous approaches, e.g. [28, 29], the Gutzwiller operator was chosen to be diagonal in the Γ-basis in order to restrict the number of variational parameters. In the limit of infinite dimensions, each of the degenerate eigenstates has the same potential energy so that a single variational parameter was used for every set of degenerate eigenvectors. The Gutzwiller variational energy can be expressed as a polynomial of the entries of the density matrix. In $d = \infty$ nonlocal entries of the density matrix appear only linearly, and the local symmetries are conserved. In general, however, the exact ground state $|\Psi_0\rangle$ of the system has different expectation values $|\langle\Gamma_i|\Psi_0\rangle|^2$. Indeed, in the diagrammatic expansion, the symmetry of the expectation values of the Γ_4 and Γ_6 configurations is broken even in 0th order. The energy differences are quite small but the deviations in the variational parameters are substantial.

In any case, we can use the spin-band symmetry in the paramagnetic case and the

spin and band symmetry in the analysis of a ferromagnetic wave function. The variational coefficients λ_{I_1,I_2} of the Gutzwiller correlator $\hat{P} = \sum_{I_1,I_2} \lambda_{I_1,I_2} |I_1\rangle\langle I_2|$ have the same structure as the matrix elements of the on-site Coulomb interaction in Eq. (3.1.19). Thus, we do exclude any matrix elements $|I_1\rangle\langle I_2|$ that include spin flips or that do not conserve the z-component of the angular momentum. For a paramagnetic $|\Psi_0\rangle$ with $E_{\mathrm{F},\uparrow} = E_{\mathrm{F},\downarrow}$, all quantities are symmetric under a change of the spin index $\uparrow, \downarrow$. Then, we can set the diagonal entries of λ_{I_1,I_2} to

$$\begin{aligned} \lambda_{u_1} &= \lambda_{d_1} = \lambda_{u_2} = \lambda_{d_2} \ , \\ \lambda_{u_1 d_1} &= \lambda_{u_2 d_2} \ , \\ \lambda_{u_1 d_2} &= \lambda_{d_1 u_2} \ , \\ \lambda_{u_1 u_2} &= \lambda_{d_1 d_2} \ , \\ \lambda_{u_1 d_1 u_2} &= \lambda_{u_1 d_1 d_2} = \lambda_{u_1 u_2 d_2} = \lambda_{d_1 u_2 d_2} \ . \end{aligned} \tag{3.1.21}$$

In the ferromagnetic case with $|\Psi_0\rangle = |\Psi_0\rangle_\uparrow \otimes |\Psi_0\rangle_\downarrow$ we can set

$$\begin{aligned} \lambda_{u_1} &= \lambda_{u_2} \ , \\ \lambda_{d_1} &= \lambda_{d_2} \ , \\ \lambda_{u_1 d_1} &= \lambda_{u_2 d_2} \ , \\ \lambda_{u_1 d_2} &= \lambda_{d_1 u_2} \ , \\ \lambda_{u_1 d_1 u_2} &= \lambda_{u_1 u_2 d_2} \ , \\ \lambda_{u_1 d_1 d_2} &= \lambda_{d_1 u_2 d_2} \ . \end{aligned} \tag{3.1.22}$$

In both cases, the remaining variational parameters are $\lambda_\varnothing$, $\lambda_{u_1 d_1 u_2 d_2}$, and the non-diagonal entries $\lambda_{u_1 d_1, u_2 d_2}$ and $\lambda_{u_1 d_2, u_2 d_1}$.

The diagonal form of the local density matrix ensures that the coefficients Z_{I_1,I_2} have the same nonvanishing entries as the coefficients λ_{I_1,I_2}, and the spin-band symmetries also apply. The same holds for the external coefficients $M_{I_1,I_2}(\hat{U})$ of the potential energy. The external coefficients $M_{I_1,I_2}(\hat{c}_{u1})$ that are used in the evaluation of the kinetic energy show a more complicated structure. The set of independent quantities is given by

$$\begin{aligned} & M^{\mathrm{kin}}_{\varnothing,u_1} \ , \\ & M^{\mathrm{kin}}_{d_1,u_1 d_1} \ , M^{\mathrm{kin}}_{d_1,u_2 d_2} \ , M^{\mathrm{kin}}_{u_2,u_2,u_1} \ , M^{\mathrm{kin}}_{d_2,d_1 u_2} \ , M^{\mathrm{kin}}_{d_2,u_1 d_2} \ , \\ & M^{\mathrm{kin}}_{u_1 d_1,u_1 u_2 d_2} \ , M^{\mathrm{kin}}_{u_1 d_2,u_1 d_1 u_2} \ , M^{\mathrm{kin}}_{d_1 u_2,u_1 d_1 u_2} \ , M^{\mathrm{kin}}_{d_1 d_2,u_1 d_1 d_2} \ , M^{\mathrm{kin}}_{u_2 d_2,u_1 u_2 d_2} \ , \\ & M^{\mathrm{kin}}_{d_1 u_2 d_2,u_1 d_1 u_2 d_2} \ . \end{aligned} \tag{3.1.23}$$

This result is valid in the paramagnetic and in the ferromagnetic case but the symmetries of the system do not allow further simplifications.

The task to build in the symmetries into the diagrammatic expansion is quite involved. Further information is given in appendix A.2.

3.1.5 Other orbital combinations

The atomic orbitals in a lattice are influenced by their neighboring atoms. In crystal-field theory the neighboring sites (called 'ligands') are considered to be monopols

that attract the electrons on its original site. This extra field splits up the $3d$ orbitals with angular momentum $l = 2$. On a square lattice, the d_{xz} and d_{yz} orbitals will be degenerate eigenstates of the local Hamilton operator that includes the ligand field. The p_x-p_y orbitals remain eigenstates because both orbitals are equally affected by the crystal field. The symmetries of the p_x-p_y and the d_{xz}-d_{yz} orbitals will lead to the same coefficients in the interaction energy and in the kinetic energy operator. Therefore, this two-band Hubbard model equally applies to p_x-p_y orbitals and to d_{xz}-d_{yz} orbitals. A more detailed discussion is deferred to appendix A.3.

3.2 Application to ferromagnetism

As seen in the previous section, the local Hamiltonian favors the formation of local magnetic moments. Eq. (3.1.20) shows that for $J > 0$ the states with maximal local spin $S = 1$ are lowest in energy in the two-particle sector, in accordance with Hund's first rule principle. Therefore, for large values of J, the ground state of the lattice system may show global ferromagnetism. Moreover, a reduction of the charge fluctuations leads to a larger probability of local states with nonvanishing magnetic moments as devised by Van Vleck ('minimum polarity model') [30]. In contrast to that, the Stoner picture gives a different explanation for the origin of ferromagnetism. In this picture, a splitting between majority and minority bands reduces their mutual Coulomb repulsion due to the Pauli principle.

In the Gutzwiller variational approach, the number of energetically costly multiple occupancies is reduced by an adjustment of the variational parameters. Therefore, we can expect that the Gutzwiller wave function predicts ferromagnetism at much larger interaction strengths than the uncorrelated SPPS. The Gutzwiller variational description leaves room both for the Stoner band splitting and the local moment formation as source for itinerant ferromagnetism.

3.2.1 Density of states

We begin our discussion with the density of states as a function of the band filling. This quantity plays an important role in the Stoner theory of ferromagnetism.

We define the quantity M as

$$\langle \hat{n}_{i\uparrow} \rangle = n_\uparrow = (1 + M)n_0 \ , \quad \langle \hat{n}_{i\uparrow} \rangle = n_\downarrow = (1 - M)n_0 \tag{3.2.1}$$

with $i \in \{1, 2\}$, so that $0 \leqslant M \leqslant 1$, and the total density remains constant. The total magnetization will be given by $M_{\text{tot}} = 2(n_\uparrow - n_\downarrow)$ when both bands are considered. In order to obtain the optimal magnetization we will perform a scan in the M-U plane while we keep the ratio of J and U fixed to $J = 8/30$. Then, we optimize the Gutzwiller energy for each magnetization as described in appendix 3.3.

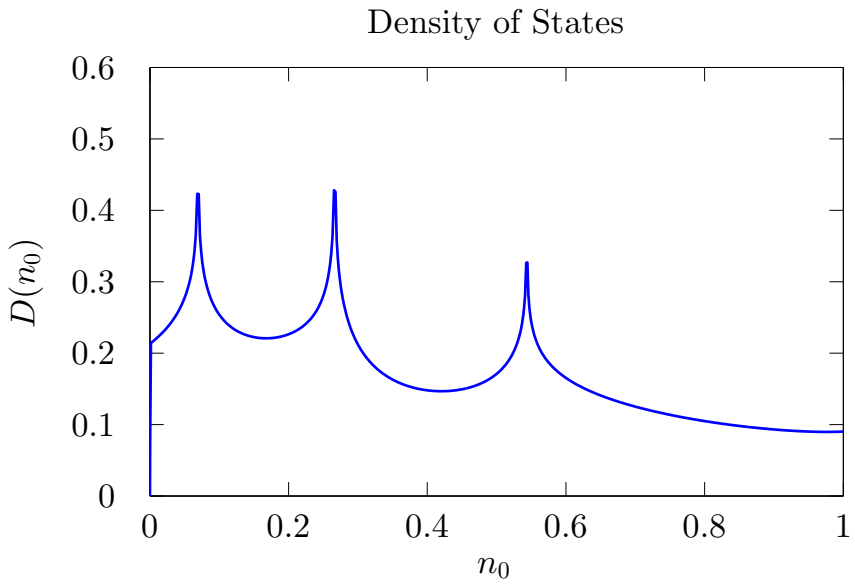


Figure 3.3: The density of states $D(n_0)$ as a function of the density n_0. The hopping amplitudes are $t_x^{11} = -1.0$, $t_y^{11} = -0.6$, $t_{xy}^{11} = +0.2$, $t_{xy}^{12} = -0.4$.

The mean-field solution of the (single-band) Hubbard model displays a finite magnetization if the interaction strength U times the density of states at the Fermi energy $D(E_\mathrm{F})$ is sufficiently large, $D(E) = \mathrm{d}n_0(E)/\,\mathrm{d}E$. The Stoner criterion for the critical interaction strength U_Stoner reads

$$U_\mathrm{Stoner}\, D(E_\mathrm{F}) \geqslant 1\ . \tag{3.2.2}$$

This simple criterion does not apply strictly in regions where the derivative of the density of states is large. Neither does it take the two-band structure of our Hamilton operator into account. However, it provides an estimate for the ferromagnetic phase transition for the HF solution if J is not too large.

Throughout this section, we focus on the kinetic energy operator with the amplitudes

$$t_x^{11} = -1.0\ ,\quad t_y^{11} = -0.6\ ,\quad t_{xy}^{11} = +0.2\ ,\quad t_{xy}^{12} = -0.4\ . \tag{3.2.3}$$

The density of states $D(n_0) = D(E)|_{E=E_\mathrm{F}(n_0)}$ in Fig. 3.3 shows three peaks at $n_0 \approx 0.069$, 0.266 and 0.543. Below we investigate the ferromagnetic transition at $n_0 = 0.2$, 0.265, 0.275, 0.3. The densities are located near the second peak in the density of states. The corresponding kinetic energies of the free system $E_\mathrm{kin}^0 = \langle\Psi_0|\hat{T}|\Psi_0\rangle$ are $E_\mathrm{kin}^0 = -1.8172$, -2.2639, -2.3274 and -2.4797, respectively. Therefore, the average kinetic energy is very similar for all cases under investigation.

3.2.2 Ferromagnetic transition

For $n_0 = 0.2$ we have $D(n_0 = 0.2) \approx 0.227$ so that the Stoner criterion gives a critical interaction strength $U_\mathrm{Stoner} \approx 4.4$. A more precise value is obtained from the Hartee-Fock approximation which corresponds to the Stoner approach in the

multi-band case but includes also the Hund's-rule coupling J. As seen in Fig. 3.4, in the Hartree-Fock approximation the magnetization jumps to a finite value at $U_{\mathrm{HF}} \approx 3.3$. The magnetization then increases monotonically until the ground-state is fully polarized at $U_{\mathrm{HF}}^{\mathrm{sat.}} \approx 3.6$. The nature of the jumps can be understood as a first-order phase transition as shown below. The Gutzwiller approach reveals a different picture. In second order, the magnetization becomes finite for $U_{\mathrm{G}} \approx 6.4$ and becomes fully magnetized for $U_{\mathrm{G}}^{\mathrm{sat.}} \approx 6.68$. This shows that the magnetization is shifted to much larger interaction strengths in the Gutzwiller wave function. The infinite-d approximation becomes magnetized at $U_{\mathrm{G}}^{\infty} \approx 6.4$ and becomes fully magnetized for $U_{\mathrm{G}}^{\infty,\mathrm{sat.}} \approx 7.35$. Therefore, the second order diagrams in our diagrammatic expansion do not change the results on ferromagnetism significantly.

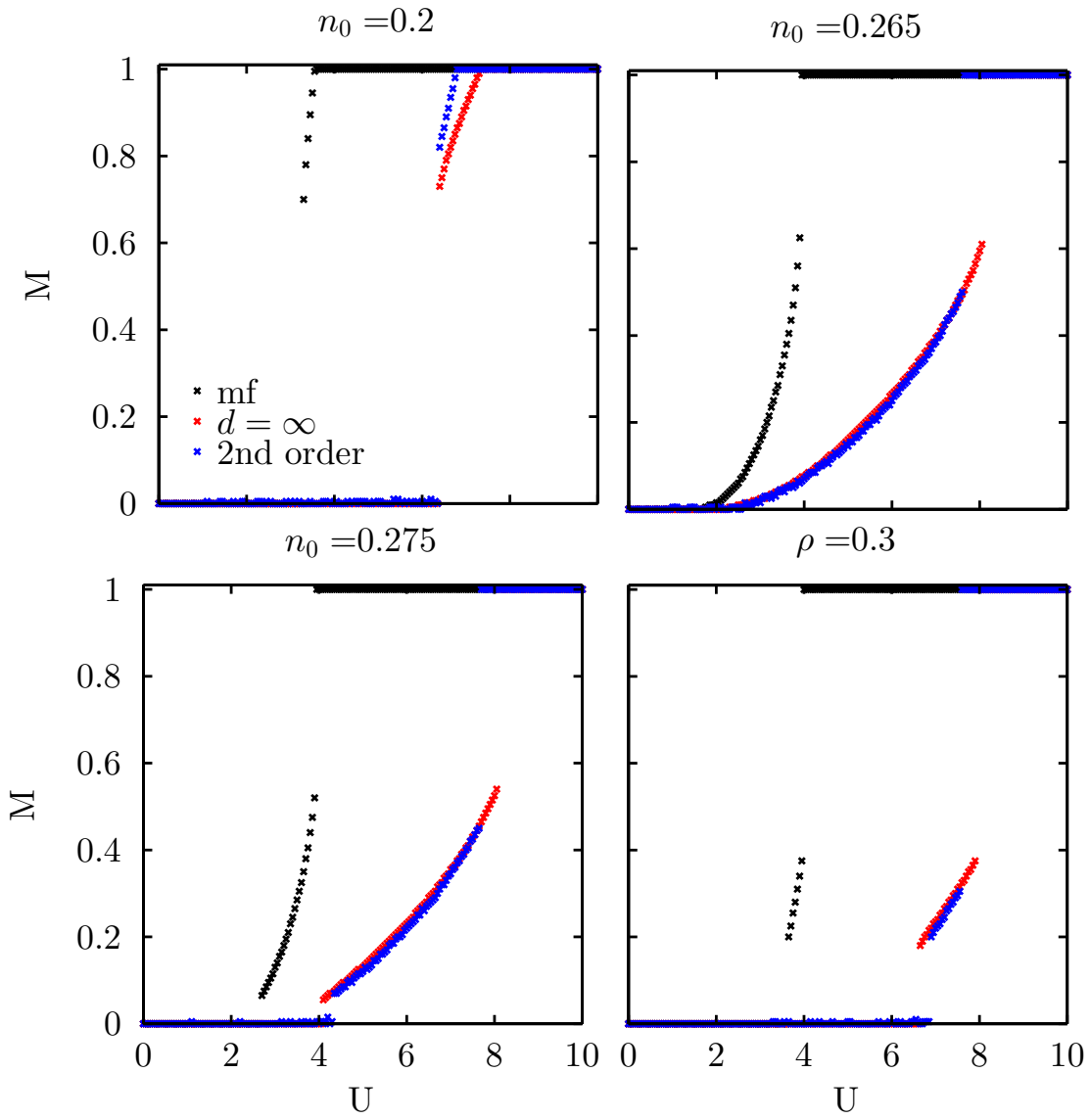


Figure 3.4: Magnetization of the Gutzwiller wave function. The black, red and blue crosses give the HF, the infinite-d and the second order approximation, respectively.

Next, we analyze the density $n_0 = 0.265$ that lies very close to the second peak in

the density of states, see Fig. 3.3. According to the Stoner criterion (3.2.2), the large density of states causes a finite magnetization at much smaller interaction strength. In our second-order Gutzwiller approach, the ground-state becomes already magnetized at $U_{\mathrm{G}} \gtrsim 2$ although a precise evaluation of the threshold is hindered by numerical difficulties. The magnetization in the second-order approximation jumps to the fully magnetized state at $U_{\mathrm{G}}^{\mathrm{sat.}} = 7.6$. The infinite-$d$ approximation lies almost on top of the second-order expansion except at the transition to the fully magnetized state which occurs at $U_{\mathrm{G}}^{\infty,\mathrm{sat.}} \approx 8.1$. The HF-result shows the same qualitative behavior but the onset of ferromagnetism is at $U_{\mathrm{HF}} < U_{\mathrm{G}}$. Moreover, the magnetization increases more rapidly as function of the interaction strengths and saturates already at $U_{\mathrm{HF}}^{\mathrm{sat.}} \approx 3.95$.

For $n_0 = 0.275$ the second-order magnetization result jumps to a finite value at $U_{\mathrm{G}} \approx 4.3$ and becomes fully spin polarized at $U_{\mathrm{G}}^{\mathrm{sat.}} \approx 7.6$. The transition points of the infinite-d (HF) approximation lie at $U_{\mathrm{G}}^{\infty} \approx 4.1$ ($U_{\mathrm{HF}} \approx 2.7$) and $U_{\mathrm{G}}^{\infty,\mathrm{sat.}} \approx 8.05$ ($U_{\mathrm{HF}}^{\mathrm{sat.}} \approx 3.95$), respectively. The magnetization curve shows the same qualitative behavior in all three approximations. The magnetization jumps to a small but finite value. Then the magnetization increases gradually as a function of U whereby the slope is much steeper in HF than in Gutzwiller theory. Lastly, the magnetization jumps to full saturation at $U^{\mathrm{sat.}}$. In general, the critical values are much larger in Gutzwiller theory than in the Hartree-Fock approach. Note that the second-order terms to the result in $d = \infty$ lead to fairly small quantitative corrections. The magnetization onset requires a larger interaction strength U for $n_0 = 0.275$ because the density of states which is lower for $n_0 = 0.275$ than for $n_0 = 0.265$. Furthermore, we can see that the transitions to the fully magnetized state occur at almost the same interaction strength as for $n_0 = 0.265$ in all approximations. This shows, that the transition to the fully polarized state depends on the density but not on the density of states.

For $n_0 = 0.3$ we still recover qualitatively the same behavior as for $n_0 = 0.275$ but the region between the onset of ferromagnetism and the transition to the fully polarized phase becomes smaller. Again, the critical values in Gutzwiller theory are about a factor of two larger than in Hartree-Fock theory.

In summary we can state that a large density of states at the Fermi energy promotes ferromagnetism. This is the essence of the Stoner criterion. The Gutzwiller approach shows, however, that ferromagnetism, in general, requires large Coulomb interactions. Moreover, the Gutzwiller approach leaves room in parameter space for non-saturated ferromagnetism.

3.2.3 Gutzwiller energies

The Gutzwiller energy E_G is shown in the left panel of Fig. 3.5 for the density $n_0 = 0.275$ and $J = 8/30U$, in comparison with the Hartree-Fock energy. As expected, the ground-state energy is much lower because the Gutzwiller wave function has much more variational freedom than the HF state. The transition to the fully polarized state shows up as a kink in the ground-state energy. On the scale of the figure,

the energies of the Gutzwiller approximation and the higher-order corrections agree very well and deviations become visible only in the vicinity of the transition points. The energy differences are of the order $\mathcal{O}(10^{-2})$, i.e., only a few percent of the total energy.

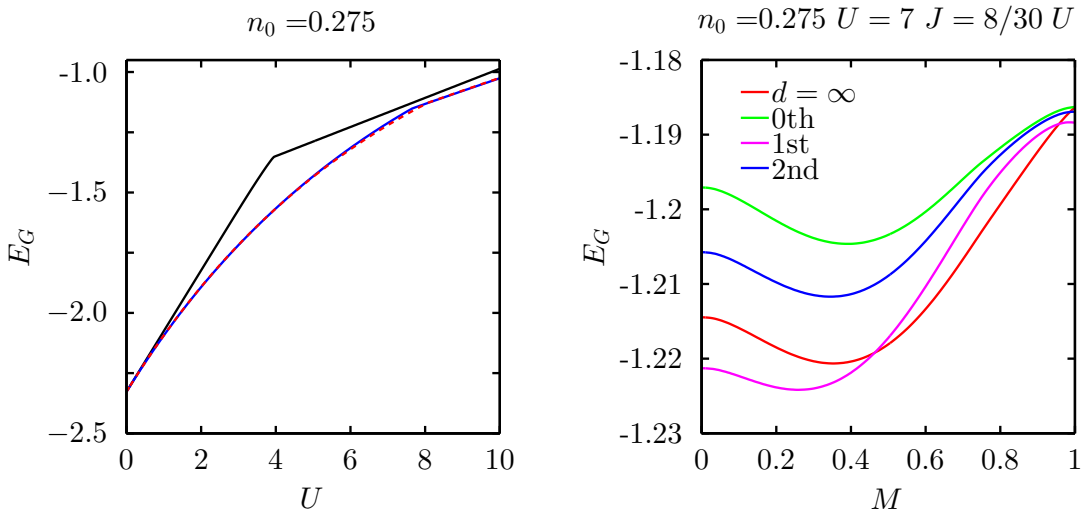


Figure 3.5: Left: Variational ground-state energy E_G as a function of the interaction strength U. The black, red and blue lines give the HF, the infinite-d and the second order approximation, respectively. Right: The optimized Gutzwiller energy E_G as a function of the magnetization and density $n_0 = 0.275$. The interaction strength is set to $U = 7.0$. The red line gives the infinite-d approximation. The green, purple and blue line give the zeroth, first and second order approximation.

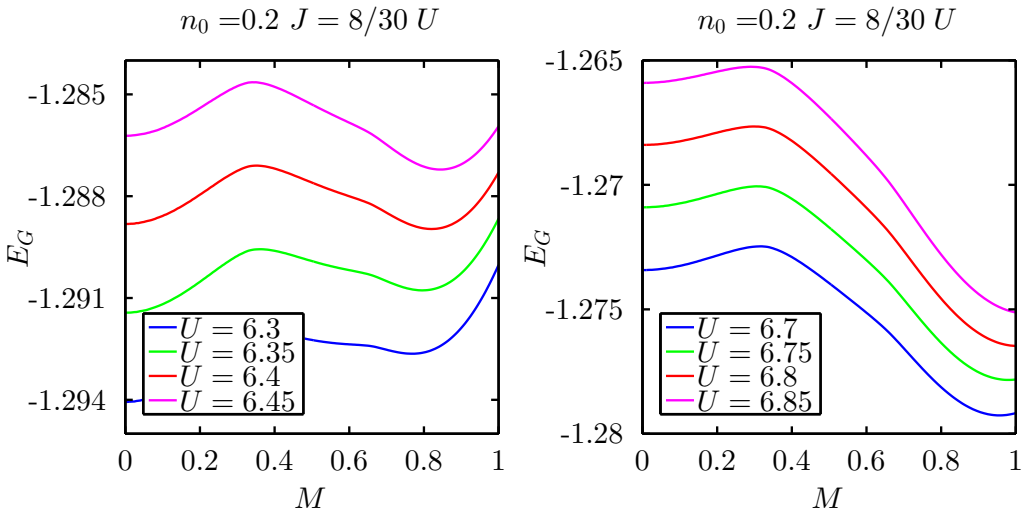


Figure 3.6: Gutzwiller energy E_G plotted as a function of the magnetization for fixed values of U for $n_0 = 0.2$. Left: The local minimum is lower than the energy for $M = 0$, if $U_{\mathrm{G}} \gtrsim 6.4$. Right: The minimum drifts towards to $M = 1$ with increasing U. The minimum vanishes at around $U_{\mathrm{G}}^{\mathrm{sat.}} = 6.8$.

The right panel of Fig. 3.5 shows the energy as a function of the magnetization for fixed interaction strength $U = 7$. The HF energy is omitted because it is much larger than the energies of the other orders. The energy differences $E_G^n - E_G^{n+1}$ between the orders give an estimate for the precision of the diagrammatic analysis. Ideally, the energy differences should be smaller than the energy differences between the maximum and minimum. However, the energy difference between the maximum and minimum of each order is as large as the energy differences $E_G^n - E_G^{n+1}$. Nevertheless, the higher-oder corrections lead primarily to a vertical shift of the magnetization curves along the energy axis. The similar shape of the curves as a function of M shows that already the infinite-d limit can be used as a good estimate for the Gutzwiller magnetization.

Next, we discuss the origin of the discontinuities in the magnetization curve. They can be understood by an examination of the energy minima. Fig. 3.6 shows the Gutzwiller energy as a function of the magnetization for the density $n_0 = 0.2$ and various interaction strengths U near the transition points. The energy has a local minimum at finite magnetization M for $U = 6.3$ as shown in the left panel. This minimum drops below the energy $E_G(M = 0)$ for $U_\mathrm{G} \approx 6.4$ so that the magnetization becomes finite. For larger values of U, the minimum moves towards $M = 1$ and vanishes vanishes at around $U_\mathrm{G}^{\mathrm{sat.}} = 6.8$, as can be seen in the right panel. In this case, the magnetization curve is continuous. For other densities, see Fig. 3.4, the transition to the fully polarized state is also discontinuous.

3.2.4 Variational parameters

The dependence of all free variational parameters on the interaction strength U is shown in Fig. 3.7. The variational parameters split up when the magnetization becomes finite. For the fully spin polarized state, the diagonal and nondiagonal parameters λ_{I_1,I_2} that contain a down-spin electron jump back to their initial values at $U = 0$. The remaining parameters $\lambda_\varnothing$, λ_{u_1} and $\lambda_{22} = \lambda_{I,I}$ with $I = u_1, u_2$ change only slightly.

When the on-site energies become too large the Gutzwiller wave function does not give a physical meaningful approximation of the ground state anymore. For large interaction strength U and densities near half-filling, some of the variational parameters vanish identically. In Fig. 3.8, we show the magnetization as a function of U for $n_0 = 0.48$ and $J = 8/30\ U$. The minimum of the energy is unique for small interaction strength. The optimization algorithm follows this minimum even for large interaction strength until $U \approx 8.5$. Then, the optimization algorithm runs into a second minimum, where the only nonvanishing free variational parameters λ_{I_1,I_2} are λ_{u_1}, $\lambda_{u_1u_1}$, $\lambda_{d_1d_2}$, $\lambda_{u_1d_1d_2}$, and $\lambda_{u_1d_1u_2d_2}$. This minimum lies on the edge of the allowed parameter space, see appendix A.2.3. The potential energy as well as the kinetic energy remain finite and the magnetization is $M \approx 0.5$. An interpolation of the energy of this minimum shows that the transition to this state will occur already at $U \gtrsim 7$. Therefore, we have to restrict the viability of our approach to the parameter regime $U \leqslant 7$. A similar behavior of the variational parameters has been found in [10], where the infinite-d limit of a similar two-band model was studied.

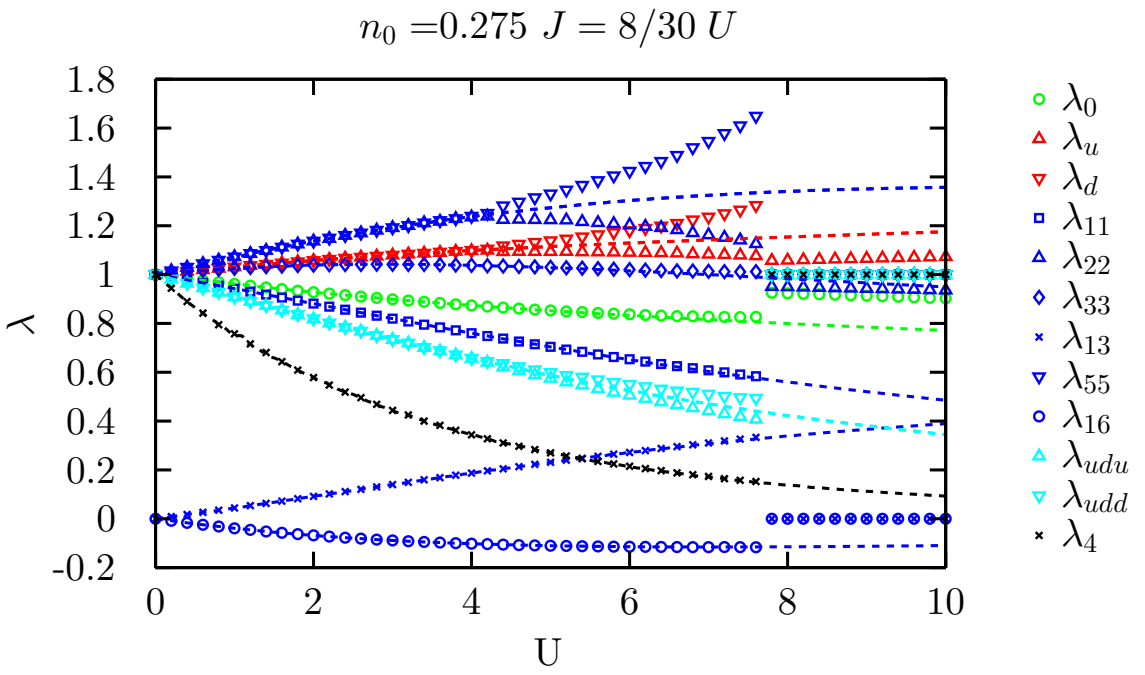


Figure 3.7: Gutzwiller variational parameters for the optimal magnetization and total density $n_0 = 0.275$. J is set to $8/30\ U$. The variational parameters λ_{I_1,I_2} with $|I_1| = 2 = |I_2|$ are labeled according to the standard ordering of the two-particle states, e.g. $1 = |\uparrow\downarrow, \varnothing\rangle$. The variational parameters split up when the magnetization becomes finite. The variational parameters that act on a state including a down-spin jump back to their initial values at $U = 0$ when $|\Psi_0\rangle$ becomes fully magnetized.

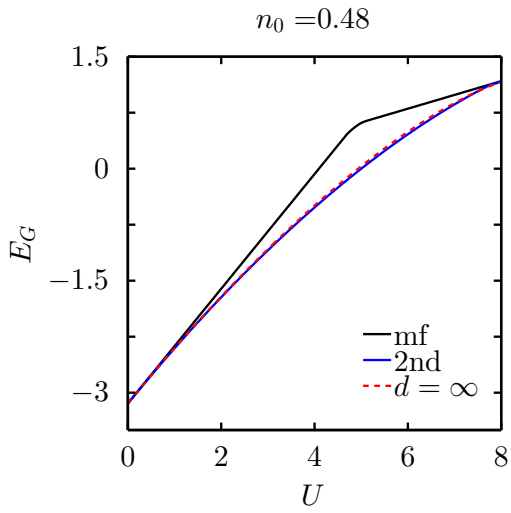


Figure 3.8: Energy plotted as a function of the interactions strength U for $n_0 = 0.48$.

3.2.5 Influence of the Hund's rule coupling

Lastly, we study the influence of the Hund's-rule coupling on the magnetization. Fig. 3.9 shows the magnetization as a function of the interaction strength J and fixed $U = 6.0$ for the densities $n_0 = 0.26$ and $n_0 = 0.48$. For $J = U/3$, the fully polarized SPPS is an eigenstate of the Hamilton operator. This eigenstate might become the ground-state if the interaction energy is chosen large enough. For values $J > U/3$ the potential energy has negative matrix elements which might lead to a BCS-type ground-state. Therefore, the interaction strength J is restricted to $J \leqslant U/3$. The ground-state is already fully magnetized in the HF-approximation for both densities.

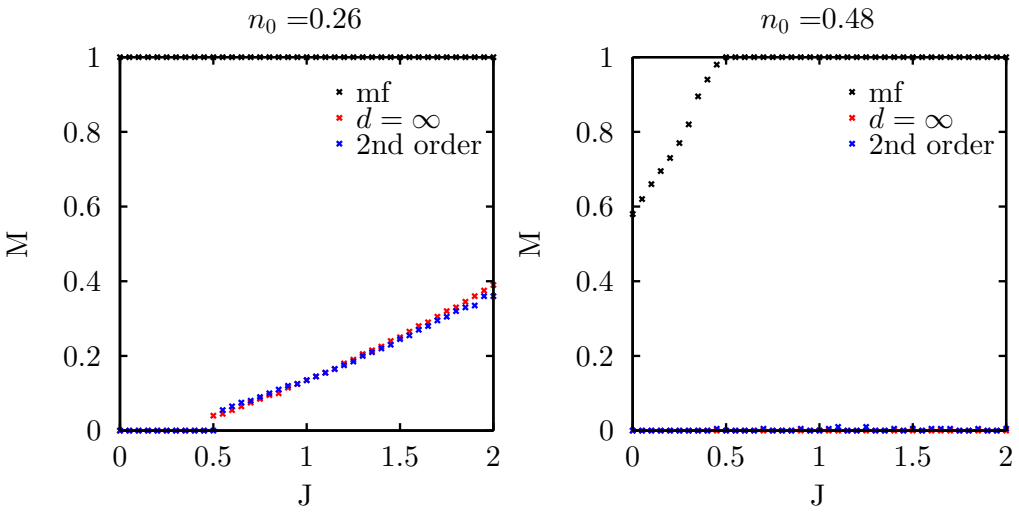


Figure 3.9: Magnetization plotted as a function of the interactions strength J for fixed $U = 6.0$ and densities $n_0 = 0.26$, $n_0 = 0.48$.

For $n_0 = 0.26$, the magnetization of the second order (infinite-d) result becomes finite at $J = 0.55$ ($J = 0.5$). Again, the second-order corrections are small. For $n_0 = 0.48$ the density of states is lower so that the magnetic phase is not visible in the Gutzwiller wave function for any $J \leqslant U/3$. In order to find a region in parameter space where the second-order corrections are more prominent, we adjust our hopping amplitudes.

In Fig. 3.10, the magnetization is shown as a function of J for a different set of parameters $t_{ij}^{\sigma\sigma'}$. The magnetization becomes finite for $J = 0.3$ in the infinite-d approximation. In second order, numerical difficulties do not permit an exact evaluation of the onset of the magnetization but it is seen clearly that the magnetization is smaller than predicted by the infinite-d approximation. For this specific setup, the density of states has a peak near $n_0 = 0.48$ so that the tendency towards ferromagnetism increases. Furthermore, the density is closer to half band-filling where the higher-order corrections become more prominent.

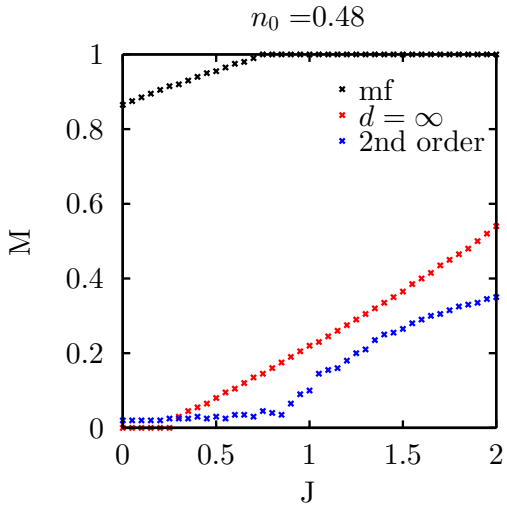


Figure 3.10: Magnetization plotted as a function of the interactions strength J for a different set of hopping amplitudes. The kinetic energy operator is given by $t_x^{11} = -1.0$, $t_y^{11} = -0.5$, $t_{xy}^{11} = 0.4$ and $t_{xy}^{12} = -0.2$. The interaction strength is set to $U = 6.0$ and the density to $n_0 = 0.48$.

3.2.6 Summary

We conclude that a large density of states at the Fermi energy promotes ferromagnetism, in accordance with the Stoner criterion. The Hunds's rule coupling J strongly enhances ferromagnetism and supports the formation of local moments. The Gutzwiller correlator enhances the weight of the two particle occupations and, therefore, increases the importance of the Hund's-rule coupling in accordance with the minimum polarity model by Van Vleck.

In general, the Gutzwiller wave function increases the critical interaction strength obtained within Hartree-Fock theory by a factor of two. Furthermore, the Gutzwiller approach broadens the parameter space in which non-saturated ferromagnetism occurs. This gives room for the experimental observations of non-saturated ferromagnetism, e.g., in transition metals such as nickel and iron. The Gutzwiller wave functions appears to be applicable for interaction strength in which the potential energy is about twice as large as the kinetic energy. Moreover, our results show that the predictions of the infinite-d approximations for ferromagnetism are reliable away from half band-filling.

3.3 Optimizing the Fermi surface

In this section, we describe the procedure to optimize the underlying single-particle product state $|\Psi_0\rangle$. The variation of the Gutzwiller energy with respect to $|\Psi_0\rangle$ leads to an effective single-particle Hamilton operator. The optimized state $|\Psi_0\rangle^*$ will be its ground state. We show that the Coulomb interactions can lead to a deformation of the Fermi surface, i.e., the Fermi surface of $|\Psi_0\rangle^*$ changes as a function of U and J. For a noticeable Fermi surface deformations, the typical ratio of the potential and

kinetic energy of the free system must be $E^0_{\text{pot}}/E^0_{\text{kin}} \approx 1.5-2$. Any successful theory that is based on the validity of the free-particle picture must incorporate the deformed Fermi surface. In Fermi-liquid theory, for example, the particle interactions lead to a renormalized dispersion relation of the excitations near the Fermi edge. However, the Fermi surface itself enters these theories as a fixed parameter and is usually chosen to be the Fermi surface of the free system with vanishing interaction potential. As seen from Gutzwiller theory, the Fermi surface of the free system is not necessarily the correct starting point.

3.3.1 Optimization algorithm

In this work, we use the optimization algorithm which was introduced in [13]. The optimization of the SPPS is carried out as a constrained optimization of the density matrix ρ_{ij}, where we merged lattice and spin-band index for notational clarity. As a first step, we express the relation between the ground state and the single-particle density matrix with the help of Lagrange parameters η_{ij}. In this way, the Gutwiller energy depends only on the density matrix and the variational parameters. The new Lagrange functional is given by

$$\begin{aligned} S_G = &\langle \hat{H} \rangle_G \left(\lambda_{I_1,I_2}, \rho_{ij}\right) - E_S \left(1 - \langle \Psi_0 | \Psi_0 \rangle\right) \\ &+ \sum_{i,j} \eta_{i,j} \left(\rho_{i,j} - \langle \Psi_0 | \hat{c}_i^\dagger \hat{c}_j | \Psi_0 \rangle \right) \ , \end{aligned} \tag{3.3.1}$$

where the normalization condition $\langle \Psi_0 | \Psi_0 \rangle = 1$ is also taken into account. The variation of S_G with respect to $|\Psi_0\rangle$ leads to an effective single-particle Hamilton operator

$$\hat{H}_{\text{eff}} = \sum_{i,j} \eta_{i,j} \hat{c}_i^\dagger \hat{c}_j \tag{3.3.2}$$

with

$$\eta_{i,j} = -\partial_{\rho_{i,j}} \langle \hat{H} \rangle_G \ . \tag{3.3.3}$$

The functional S_G will become extremal if $|\Psi_0\rangle$ is an eigenstate of the effective Hamilton operator,

$$\hat{H}_{\text{eff}} |\Psi_0\rangle = E_S |\Psi_0\rangle \ . \tag{3.3.4}$$

In order to solve these equations self-consistently, we usually start with the ground state $|\Psi_0\rangle^{(0)}$ of the free system. We compute the density matrix and compute the coefficients η_{ij}. Then, we define the ground state of the effective Hamilton operator as the new SPPS $|\Psi_0\rangle^{(n)}$. Now, we construct a new density matrix $\rho^{(n)}$ and compute the next set of coefficients $\eta_{ij}^{(n)}$. The optimization terminates if $\Delta\eta = \eta_{ij}^{(n)} - \eta_{ij}^{(n-1)}$ drops below some threshold.

In some cases the iteration procedure can show bad convergence or lead to alternating jumps between two points. In these cases it is useful to set $\tilde{\eta}_{ij}^{(n)} = \beta\eta_{ij}^{(n)} + (1-\beta)\eta_{ij}^{(n-1)}$ as in [13] so that the iteration will smooth out. A good

choice seems to be $\beta \approx 0.7 - 0.9$. In our calculations, we usually find a fix-point of this mapping after $n < 15$ iteration steps for a threshold $\Delta\eta_{\min} = \beta * 10^{-3}$. In order to test the stability of the algorithm, we can start with a different initial state. This initial state can be constructed from a perturbed kinetic energy operator $\hat{T} + \delta\hat{t}$. Usually the optimization algorithm remains stable against these perturbations but in some cases the fix-point of this map does not need to be unique, as we shown in the section 3.3.4. For $U = 0 = J$ the initial state is the fix-point of this iteration. The gradient in Eq. (3.3.3) always incorporates all lines that are used in the diagrammatic expansion. The calculations presented in this section include all lines $\rho_{ij}^{\sigma\sigma'}$ with dmax $= ||i - j||_1 \leqslant 5$, where $||.||_1$ gives the one-norm (Manhattan metric) of the displacement vector $i - j$.

Note that the symmetries of the spin-band symmetric system fix the local entries of the density matrix. Furthermore, the kinetic energy is only rescaled by a simple scalar factor in the infinite-d approximation. Therefore, an optimization in the infinite-d approximation will not change $|\Psi_0\rangle$ in the paramagnetic case. All results in this section can, therefore, be understood as effect of the expansion to higher orders.

3.3.2 Fermi surface deformations

In the following setup the hopping amplitudes are set to

$$t_x^{11} = -1.0\,, \quad t_y^{11} = -0.6\,, \quad t_{xy}^{11} = +0.2\,, \quad t_{xy}^{12} = -0.4\,, \tag{3.3.5}$$

which is the same parameter set used in section 3.2. The equipotential lines of the lower and upper energy bands $\epsilon^{\pm}(k)$ are shown in Fig. 3.11 in order to give an overview of the Fermi surfaces for various Fermi energies. The density of states, shown in Fig 3.3, is smooth for all densities investigated in this subsection, $0.45 \leqslant n_0 \leqslant 0.52$.

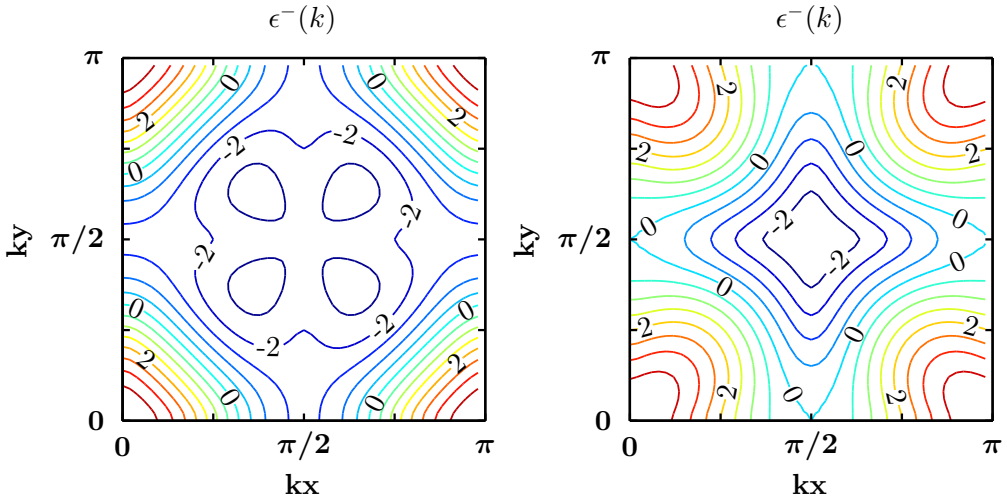


Figure 3.11: Equipotential lines of the kinetic energy for the lower (left) and upper band (right). The energy difference between two lines is set to $\Delta E = 0.5$. The Fermi surfaces display the symmetries of the square lattice.

The optimization of the SPPS leads to a deformation of its Fermi surface. In Fig. 3.12, the Fermi edges of the initial SPPS $|\Psi_0\rangle^{(0)}$ and optimized SPPS $|\Psi_0\rangle^*$ are shown for the densities $n_0 = 0.45, 0.48, 0.5, 0.52$. The Hubbard and Hund parameters are set to $U = 6.0$ and $J = 0.8$. The Gutzwiller wave function stays paramagnetic for all of these parameters as can be seen in Fig. 3.9. The Fermi surface starts to deform at $n_0 \approx 0.45$, and the deformations become more prominent for higher densities.

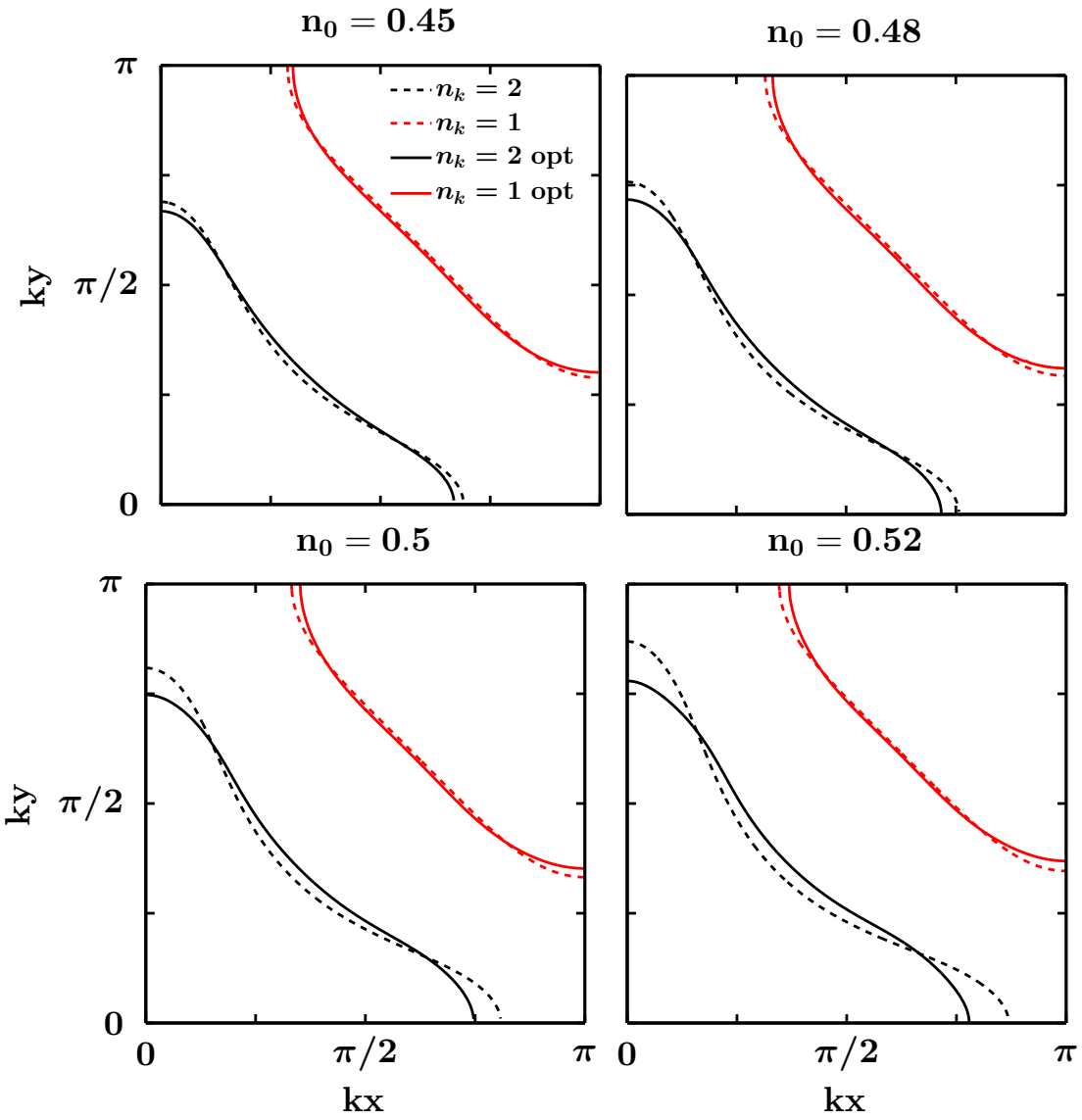


Figure 3.12: Fermi surface deformations for densities $n_0 = 0.45$, 0.48, 0.5 and 0.52. The local interaction strengths are set to $U = 6.0$ and $J = 0.8$. The dashed lines give the initial Fermi edge and the solid lines give the optimized Fermi surface. Both bands are occupied in the region between the origin and the solid (dashed) black line. In the region between the black and the red lines only the lower band is occupied.

The outer Fermi edge between $\langle \hat{n}_k \rangle = n_k = 1$ and $n_k = 0$ appears to be more robust. In our model, the infinite-d approximation does not change the Fermi surface of the

free system but the higher-order corrections lead to visible deformations of the Fermi surface. In contrast to Fermi liquid theory, the deformations of the Fermi surface are small but macroscopic. Numerical difficulties do not allow the computation of densities much higher than half-filling. However, particle-hole symmetry, described in appendix A.3.4, can be used to obtain results for $n_0 \geqslant 0.5$ by reversing the sign of t_{xy}^{11}. In this way we can show that the deformations vanish for densities $n_0 > 0.6$. For a further discussion, see appendix A.3.4.

The values for the optimized variational parameters are listed exemplarily for the parameters $n_0 = 0.48$ and $U = 6.0$ and $J = 0.8$ in table 3.1.

$\lambda_{\varnothing}$	0.39	λ_{u_1}	0.93
$\lambda_{u_1 d_1}$	0.92	$\lambda_{u_1 u_2}$	1.49
$\lambda_{u_1 d_2}$	1.26	$\lambda_{u_1 d_2, d_1 u_2}$	0.23
$\lambda_{u_1 d_1, u_2 d_2}$	−0.12	$\lambda_{u_1 d_1 u_2}$	0.83
$\lambda_{u_1 d_1 u_2 d_2}$	0.38		

Table 3.1: Optimized variational parameters for $n_0 = 0.48$, $U = 6.0$ and $J = 0.8$

The values of the coefficients λ_{I_1,I_2} indicate an increase of the weight of the energetically favorable states given in subsection 3.1.3. In accordance with to Hund's rule, the weight of the states with the highest spin multiplicity is increased.

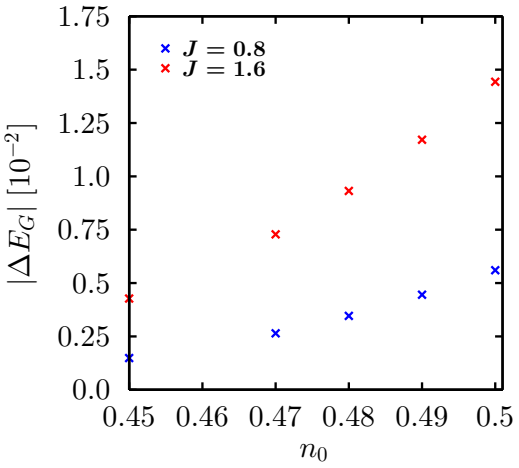


Figure 3.13: Energy gain after the optimization of $|\Psi_0\rangle$ for $U = 0.6$ is given as a function of the density n_0. The blue and red crosses give the energy difference for $J = 0.8$ and $J = 1.6$ respectively.

Typically, the value of the variational parameter changes only by a few percent in the optimization procedure of the SPPS. This indicates that the two states are very close in energy. The energy difference ΔE_G between the Gutzwiller wave function with optimized SPPS and with initial SPPS is shown in Fig. 3.13. Here, we fix $U = 6.0$ and vary the density n_0 for $J = 0.8$ and $J = 1.6$. The energy gain seems to

increase almost linearly as a function of n_0. The energy gain for $J = 1.6$ is about three times larger than the energy gain for $J = 0.8$. The energy differences are of the order of one percent of the ground-state energy, i.e., in the range of thermal energies at room temperature if we assume that our energies are given in electron Volts.

Note that the total energy gain is smaller than the typical energy differences between the different orders. In addition, the difference of the Gutzwiller energy for different cutoff lengths dmax $= 4, 5$ are of order $\mathcal{O}(10^{-2})$. However, we assume that the higher-order contributions will not change the results qualitatively, i.e., the higher orders will change the Fermi surface deformations quantitatively but the deformation itself can be considered as a feature of the Gutzwiller wave function. This point of view is supported by an analysis of the higher-order contributions in the single-band case, see subsection 3.3.4. Furthermore, we construct the initial state from a kinetic energy operator where a perturbation has been added to the hopping amplitudes. We tested the optimization algorithm for the perturbations $\delta^{22}_{xx} = \pm 0.2, \pm 0.4$, $\delta^{11}_{xy} = \pm 0.2, \pm 0.4$ and $\delta^{11}_{xy} = \pm 0.1, \pm 0.2$. The optimization algorithm converges to the same result in all cases. This shows that the distorted Fermi surface is a stable feature of the Gutzwiller wave function.

In the left panel of Fig. 3.14, we show the dependence of the Fermi surface deformations on the interaction strength J for fixed $n_0 = 0.48$ and $U = 6.0$. The size of the deformations and the energy gain ΔE_G increase with J. The Fermi surface deformations and the energy gain ΔE_G are shown for the interaction strength $U = 3$, 4, 5 and 6 in Fig. 3.15 for fixed $n_0 = 0.48$ and $J = 0.8$. It is seen that the strength of the deformations and the energy gain ΔE_G increase with U. Note that the energy scale is of the order of $10 - 100$°K if we assume that our energies are given in electron Volts. The energy gain increases gradually so that no fine tuning of the parameters is necessary to observe these effects.

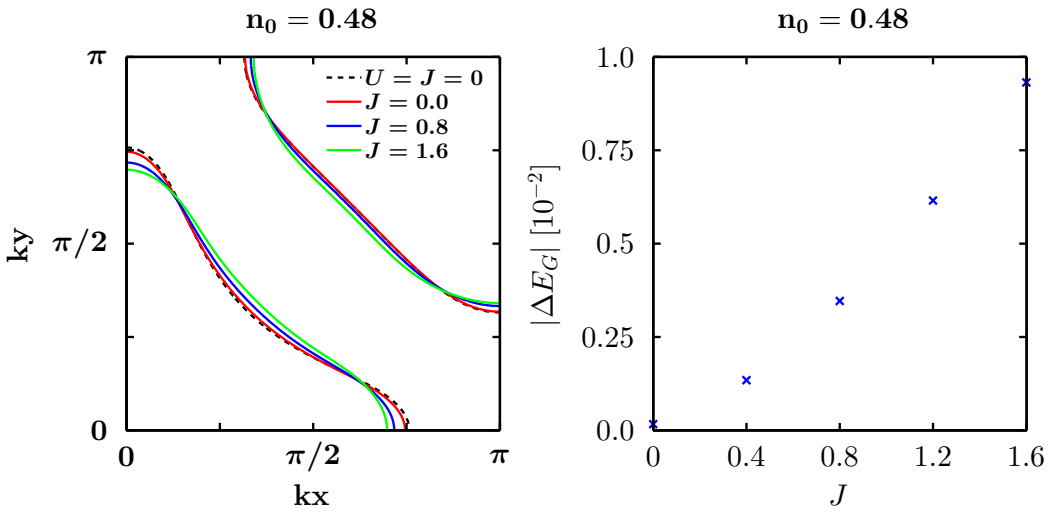


Figure 3.14: Left: Fermi surface deformations for different interaction strength J. The density and the interaction strength are set to $n_0 = 0.48$ and $U = 6.0$, respectively. The deformations increase with J. Right: Energy gain after the optimization of the Fermi surface.

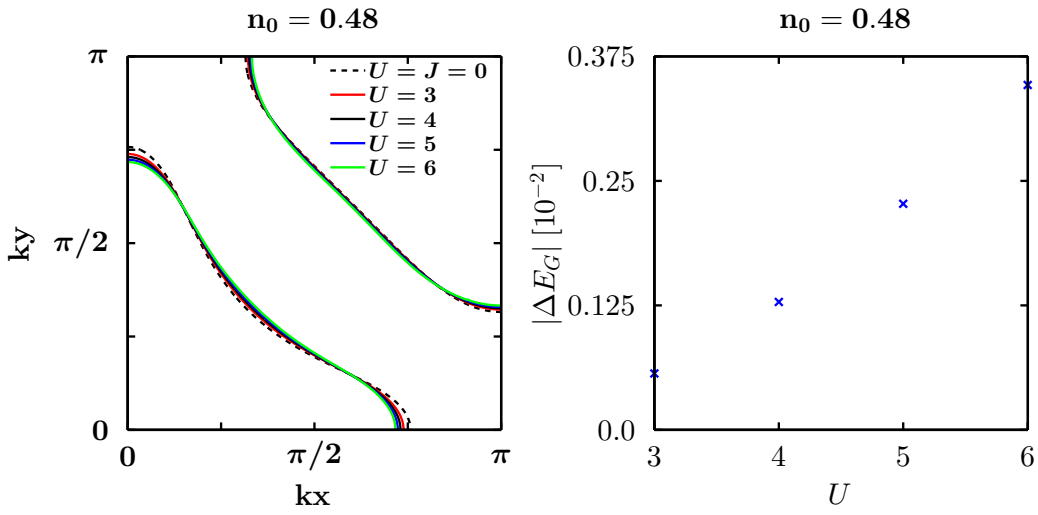


Figure 3.15: Left: Fermi surface deformations for different interaction strength U. The density and the interaction strength are set to $n_0 = 0.48$ and $J = 0.8$, respectively. The deformations increase with U. Right: Energy gain after the optimization of the Fermi surface.

3.3.3 Changing the Fermi surface topology

In the last subsection we have already seen that the optimization of the SPPS can lead to a deformed Fermi surface. In this subsection, we show that the Fermi surface can also change its topology. This effect has already been studied in [25] for the single-band Hubbard model with a Gutzwiller-Jastrow variational approach.

We examine the Fermi surface deformations for the following parameter set,

$$t_x^{11} = -1.0\,, \quad t_y^{11} = -0.5\,, \quad t_{xy}^{11} = 0.4\,, \quad t_{xy}^{12} = -0.2\,, \quad U = 6.0\,, \quad J = 0.8\,. \tag{3.3.6}$$

The amplitudes are chosen in a way such that the topology of the Fermi surface changes near half-filling where the effect of the Gutzwiller correlator is strongest.

The density of states is shown in Fig. 3.16. The peak near $n_0 = 0.47$ is caused by the change in the topology of the Fermi surface. The equipotential lines of the lower and upper energy bands $\epsilon^{\pm}(k)$ are shown in Fig. 3.17 and give an overview of the Fermi surface structure.

Again, the optimization procedure includes all lines $\rho_{ij}^{\sigma\sigma'}$ with dmax $= ||i-j||_1 \leqslant 5$. In some cases, the optimization algorithm alternates between two fix points which are energetically very close. However, the Fermi surface of these fix points may differ significantly. For example, the energetically higher fix-point can show additional islands in the Brillouin zone in which both bands are occupied. In these cases, the fix point with lower energy must be selected by hand.

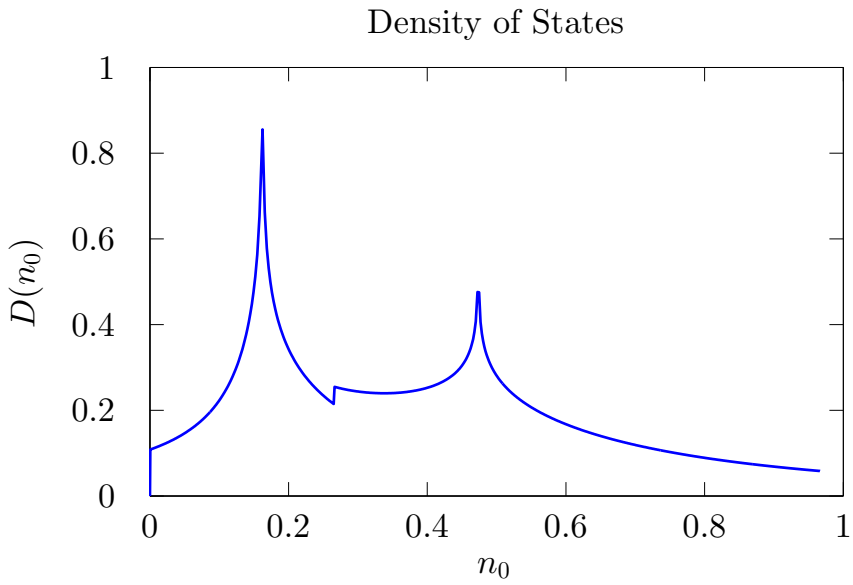


Figure 3.16: The density of states $D(n_0)$ as a function of the density n_0. The hopping amplitudes are $t_x^{11} = -1.0$, $t_y^{11} = -0.5$, $t_{xy}^{11} = +0.4$, $t_{xy}^{12} = -0.2$.

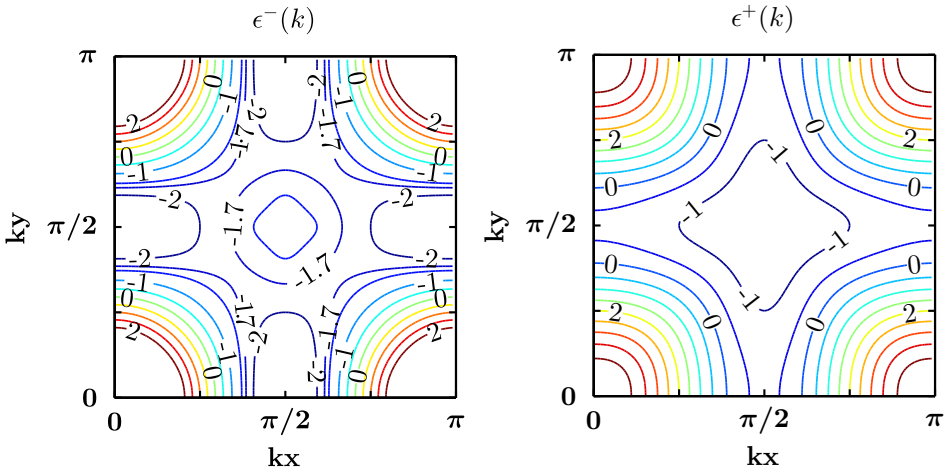


Figure 3.17: Equipotential lines of the kinetic energy for the lower (left) and upper band (right). The energy difference between two lines is set to $\Delta E = 0.5$. Additionally, the equipotential line for $\epsilon^-(k) = -1.7$ is shown for the lower band.

In Fig. 3.18, the Fermi edges of the initial and optimized SPPS are shown for the densities $n_0 = 0.4$, 0.45, 0.48, 0.50, 0.52 and 0.53. The Hubbard/Hund parameters are set to $U = 6.0$ and $J = 0.8$. Although the SPPS $|\Psi_0\rangle$ can have a small but finite magnetization in this parameter regime, as shown in Fig. 3.10, we restrict ourselves to a paramagnetic wave function. The deformation of the inner Fermi surface between $n_k = 2$ and $n_k = 1$ start for densities $n_0 > 0.4$. The outer Fermi edge between $n_k = 1$ and $n_k = 0$ is more robust. For densities $n_0 > 0.48$, the optimized inner Fermi edge still has a closed topology while the initial Fermi surface topology is open. The optimization becomes difficult for densities larger than $n_0 = 0.53$. Alternatively, the particle-hole symmetry can be used to determine the optimal Gutzwiller wave function. In this way, we can show that the deviations in the Fermi

surface are small for $n_0 \gtrsim 0.6$ where the topology of the optimized Fermi surface becomes open.

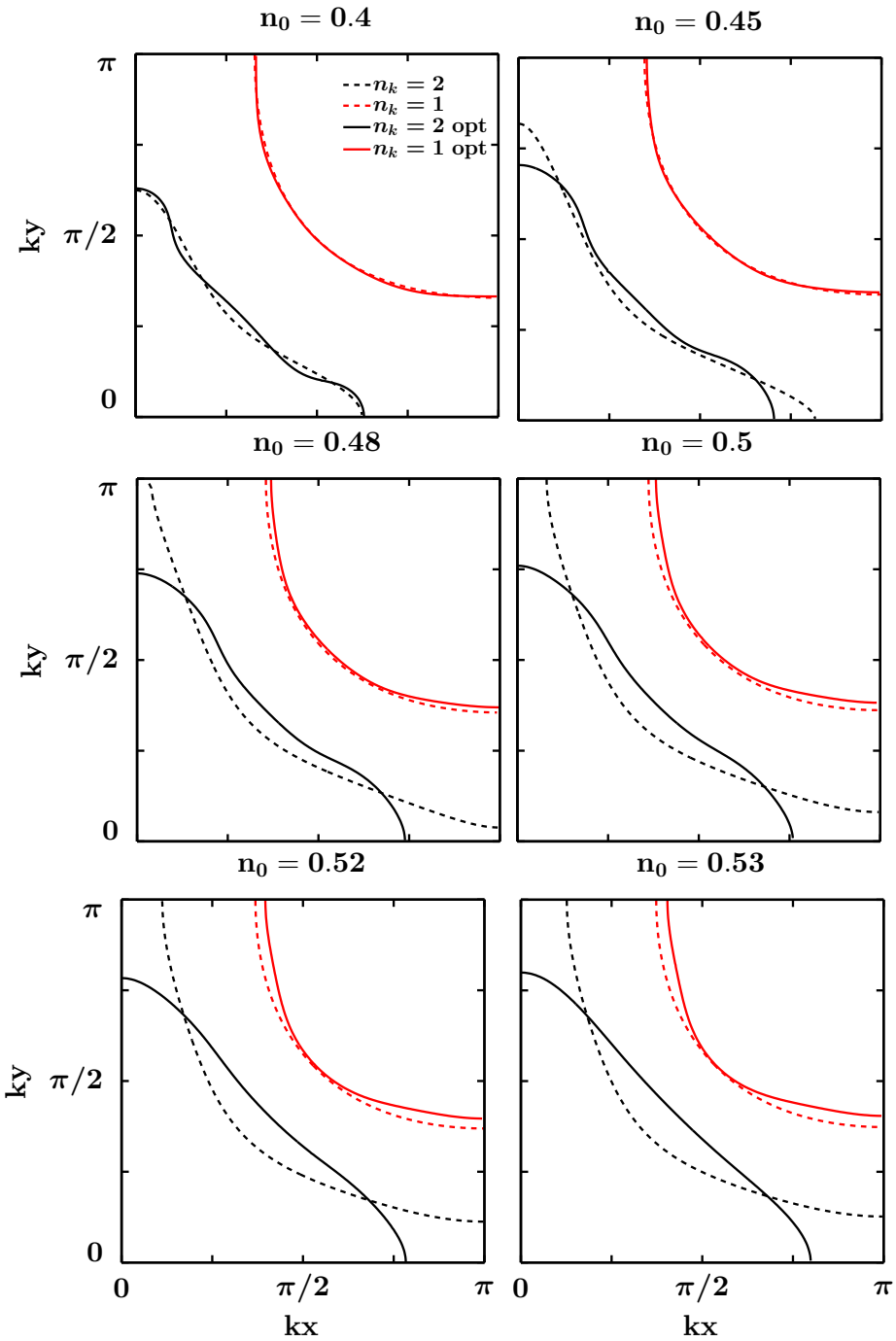


Figure 3.18: Fermi surface deformations for densities $n_0 = 0.4$, 0.45, 0.48, 0.50, 0.52 and 0.53. The local interaction strengths are set to $U = 6.0$ and $J = 0.8$. The dashed lines give the initial Fermi edge and the solid lines give the optimized Fermi surface. Both bands are occupied in the region between the origin and the solid (dashed) black line. In the region between the black and the red lines only the lower band is occupied. For densities $n_0 > 0.48$, the optimized inner Fermi edge has a closed topology while the initial Fermi surface topology is open.

The dependence of the Fermi surface deformations on the Hund's-rule coupling J is shown in Fig. 3.19. The Fermi edge for $n_k = 2$ remains open for vanishing J. An increase of the Hund's-rule interaction strength to $J = 0.4$ leads to the appearance of small islands in which both bands are filled. These islands collapse when we further increase the interaction strength to $J = 0.8$ so that the Fermi surface becomes closed. The right panel of Fig. 3.19 shows that the energy gain ΔE increases linearly in J and becomes vanishingly small for $J = 0$. From an energetic point of view, the transition from an open to a closed inner Fermi surface is gradual as a function of J. The existence of intermediate islands also shows that the hopping matrix elements gradually change.

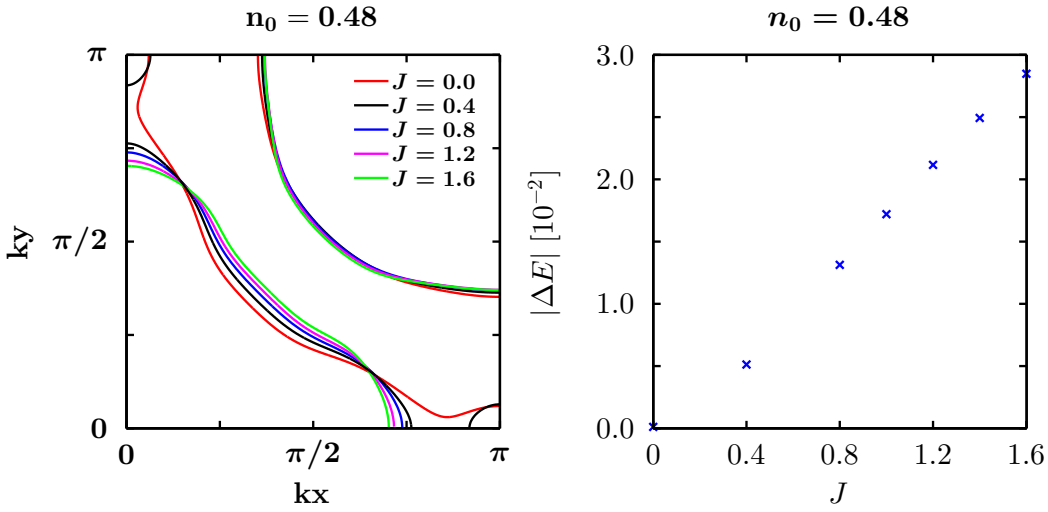


Figure 3.19: Left: Fermi surface deformations for different interaction strength J. The density and the interaction strength are set to $n_0 = 0.48$ and $U = 6.0$, respectively. The deformations increase for larger values of J. For $J = 0$, the Fermi surface topology (for $n_k = 2$) is still open. Right: The energy gain ΔE increases linearly in J.

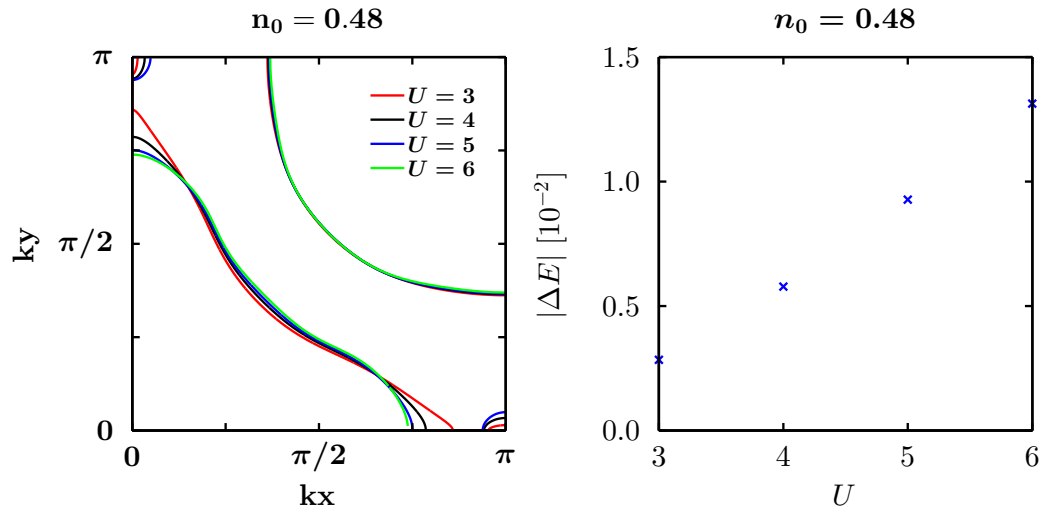


Figure 3.20: Fermi surface deformation for different interaction strength U. The density and the Hund's-rule coupling are set to $n_0 = 0.48$ and $J = 0.8$ respectively. Small islands appear for $U = 3$ in which $n_k = 2$. Right: Energy gain ΔE as a function of U.

The change in the Fermi-surface topology as a function of U for $n_0 = 0.48$ and $J = 0.8$ is shown in Fig. 3.20. For an interaction strength $U = 3$, the Fermi surface starts to deform from an open to a closed topology and small islands appear. The islands at the border of the Brillouin zone vanish for $U = 6$ again. The energy gain ΔE increases linearly in U as shown in the right panel of Fig. 3.20.

3.3.4 Pomeranchuk instability

In this section we discuss the asymmetric deformation of the Fermi surface for two independent bands with local Hubbard interaction. Within the Gutzwiller approach, this so-called Pomeranchuk instability [31] was first analyzed in [13]. Some results obtained with a mean-field theory and a renormalization group approach can be found in [32–34].

In order to test our method against previous results for the single-band Hubbard model, we set $J = 0$ and use the bare Hubbard interaction,

$$\hat{U} = U(\hat{n}_{1\uparrow}\hat{n}_{1\downarrow} + \hat{n}_{2\uparrow}\hat{n}_{2\downarrow}) \,, \tag{3.3.7}$$

with $U = 10$. The hopping amplitudes are set to

$$t_x^{11} = -1.0 \,, \quad t_y^{11} = -1.0 \,, \quad t_{xy}^{11} = 0.25 \,, \quad t_{xy}^{12} = 0.0 \,. \tag{3.3.8}$$

The density of states is smooth at half filling as can be seen in Fig 3.21.

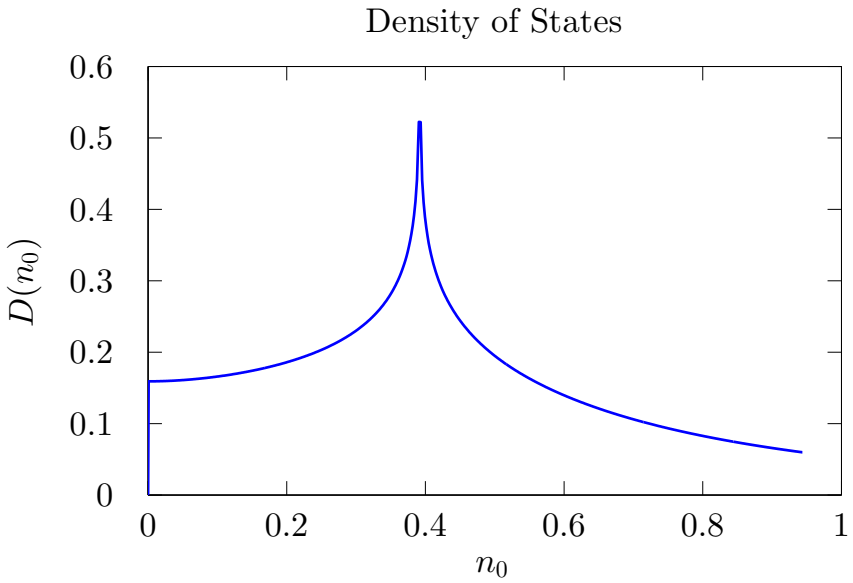


Figure 3.21: The density of states $D(n_0)$ as a function of the density n_0. The hopping amplitudes are $t_x^{11} = -1.0$, $t_y^{11} = -1.0$, $t_{xy}^{11} = 0.25$, $t_{xy}^{12} = 0.0$.

The Gutzwiller operator separates into

$$\hat{P} = \hat{P}^1 \otimes \hat{P}^2 \,. \tag{3.3.9}$$

For this special form of the Gutzwiller operator, all internal nodes Z_{I_1,I_2} with $|I_1| > 2$ vanish automatically. The same holds for $M_I(\hat{U})$ with $|I| > 4$ and for all $M_{I_1,I_2}(\hat{c}_\sigma)$ with $|I_1| + |I_2| > 3$. Furthermore, the index sets of the nonvanishing coefficients do not contains indices of the first and the second band simultaneously.

To make progress, we add a perturbation $\delta^{11}_{yy} = \delta^{22}_{xx} = 0.2$ to the kinetic energy operator in order to introduce a spatial anistropy into our system. Then, we use the ground state of this operator as the initial SPPS $|\Psi_0\rangle$. In the left panel of Fig. 3.22, we compare the optimized SPPS with and without such a perturbation. Without perturbation, both bands lie on top of each other. In the symmetry-broken case, the Fermi edge of the first band can be transformed to the Fermi edge of the second band by a rotation of $\pi/2$. The initial anisotropy, introduced by δ^{11}_{yy}, is enhanced after the optimization of $|\Psi_0\rangle$. This observation implies that the system accepts (and even enhances) the offered symmetry breaking.

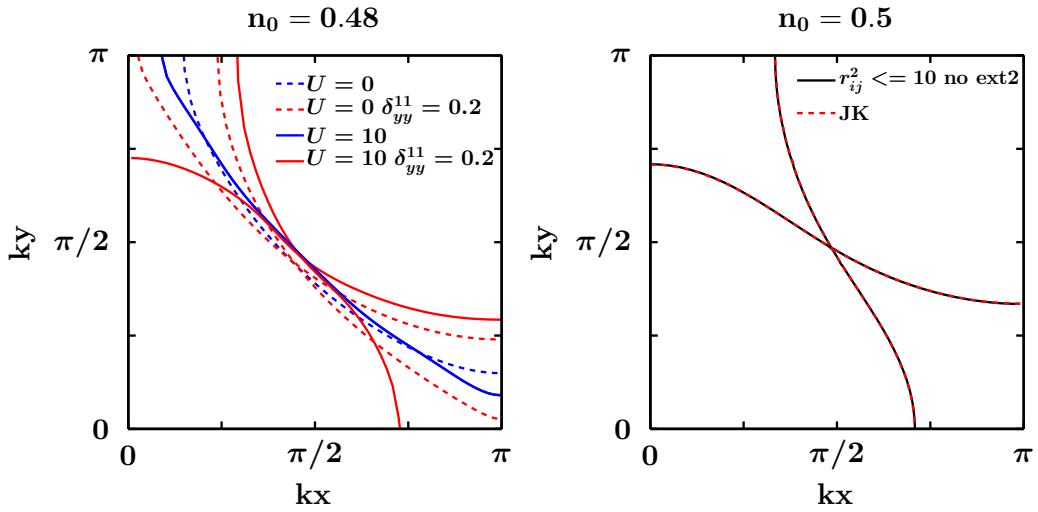


Figure 3.22: Fermi surfaces for $U = 10$. Left: The optimized Fermi surface with and without perturbation are given by the solid red and blue lines, respectively. The dashed lines give the initial Fermi surface for $n_0 = 0.48$. The initial symmetry breaking introduced by δ^{11}_{yy} is enhanced after the optimization. Right: Fermi surface to second-order, compared with the second-order result obtained by Jan Kaczmarczyk at $U = 10$ and $n_0 = 0.5$.

For the single-band Hubbard model, a higher order expansion of the diagrammatic expansion was carried out by Jan Kaczmarczyk et. al. in [35]. We compare our results with the results of an expansion up to 6th order by Jan Kaczmarczyk [36]. The cutoff parameter in the diagrammatic expansion is defined as $\rho_{ij} = 0 \; \forall r^2_{ij} > 10$. Here, $r_{ij} = ||i - j||_2$ gives the Euclidean norm of the displacement vector $i - j$. These calculations use the calculation scheme in which the external nodes with two lines are expressed by the value of the external nodes with four lines as discussed in A.4.1. In the single-band case, this guarantees that $\langle \hat{n}_\sigma \rangle = n_0$. The weight of all external nodes with two lines vanishes automatically at half filling. In this way, some long-ranged diagrams cancel out, so that the convergence with respect to the cut-off parameter dmax is improved. Our second-order results agree perfectly with the calculations obtained by [36] if we reduce the density matrix to these values as

is shown in the right panel of Fig. 3.22.

At half filling, the data in [13] agree perfectly with the fourth order expansion in [36]. However, the authors in [13] elude that all Fermi edges for $n_0 \in [0.375, 0.475]$ fall on top of each other for $k_y = 0$. We do not see this behavior in our second order data (which does not include the computation scheme in A.4.1 for $n_0 < 0.5$), so that a direct comparison of a 4th order expansion away from half filling would be necessary to clarify this issue.

The higher-orders expansion by [36] is shown in the left panel of Fig. 3.23. The deviations between the orders are still large but the fifth and sixth order agree very well. The energy gain ΔE is shown in the right panel of Fig. 3.23 as a function of the density n_0. Strong fluctuations between the results of the different orders are visible. Nevertheless, the Pomeranchuk instability is present even in third order when the energy gain is lowest. Therefore, we are confident that the Pomeranchuk instability is a true feature of the Gutzwiller wave function.

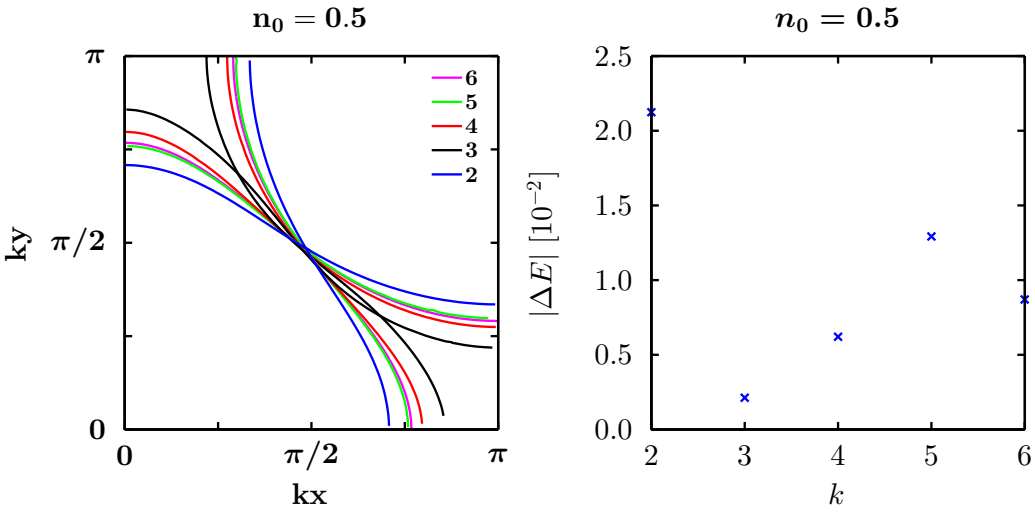


Figure 3.23: Results of the diagrammatic expansion to higher orders. Left: The optimization of the Fermi surfaces will lead to a broken symmetry in all orders. Right: The energy gain is plotted as a function of the order k. The convergence properties of the energy gain are still unsatisfactory.

3.4 Summary

In this section, we showed that the second order corrections lead to interaction-induced deformations of the Fermi surface. In contrast to the ferromagentic case, the local densities do not change due to the symmetries of the system. Therefore, the higher-order gain significance compared to the infinite-d approximation. The Fermi surface deformations are clearly visible and, therefore, do not match the assumptions made in Fermi liquid theory that the Fermi surface of the non-interacting electrons is identical to the quasi-particle Fermi surface. Moreover, we showed that the even the topology of the Fermi surface may changes as a function of the Coulomb

interaction-strength. In both cases, the energy gain after the Fermi surface optimization becomes vanishingly small for vanishing Hund's-rule coupling J. A higher order expansion of the single-band Hubbard model showed that we can safely assume that the Fermi surface deformations are true features of the Gutzwiller wave function. Moreover, we showed that the Fermi surface deformations can break the lattice symmetry in this model. The energy gain after the Fermi surface optimization is of the same order as thermal fluctuation at room temperature. Therefore, a distortion of a lattice can be induced by the correlation effects of the electrons.

Chapter 4

Conclusions and outlook

4.1 Summary

In this work, we formulated and applied the Gutzwiller variational many-body approach to multi-band Hubbard models.

In chapter 1, we gave a short introduction to the problem and an outline of the scope of the work. In the chapter 2, we developed a complete, concise diagrammatic formalism for a perturbative evaluation of expectation values for Gutzwiller-correlated wave functions on finite lattices. The derivation of the diagrammatic expansion consists of three steps.

In a first step, we introduced a one-to-one mapping between a sequence of fermion operators and their Hartree–Fock counterparts in order to eliminate all local contractions. We explicitly showed the consistency of the mapping. In a second step, we derived and applied the linked-cluster theorem. To this end, we expanded numerator and denominator in the Gutzwiller expectation value of one-site and two-site operators in terms of a perturbation series, and used Wick's theorem to express the coefficients in terms of diagrams. The introduction of the Hartree-Fock operators excludes all local contractions so that lines between identical lattice sites are zero by definition. The normal ordering of the operators and the sum over distinctive lattice sites permitted the introduction of Graßmann variables. For multi-band Gutzwiller wave functions, we had to introduce a formal representation of local operators in terms of an exponential series which led to a re-definition of the values of external and internal vertices. Then, the linked-cluster theorem applied, both for infinite *and* finite lattices, i.e., the unconnected diagrams in the numerator are canceled by the denominator. In this way, the nth-order in perturbation theory corresponds to summing all connected diagrams with n internal nodes. As a third and last step, we eliminated all internal nodes with two lines by fixing a subset of our variational parameters. We showed that, for our applications, this gauge fixing does not restrict the variational freedom of the Gutzwiller correlator. Furthermore, we discussed the implications of this parameter gauge for a more general setup.

The big advantage of our diagrammatic approach lies in the fact that it simplifies

decisively in the limit of infinite spatial dimensions. We obtain the exact result for Gutzwiller expectation values of single-site and two-site operators without calculating a single, non-trivial diagram. Of course, the diagrams with internal vertices contribute in finite dimensions, and their importance for phase transitions and the Fermi surface has to be studied.

In the chapter 3, we therefore investigated a two-band Hubbard model on a square lattice. We introduced the Hamiltonian for two degenerate p_x-p_y (or d_{xy}-d_{yz}) orbitals where we considered electron transfers between nearest neighbors and next-nearest neighbors. The orbital degeneracy reduces the number of different hopping parameters but transitions between the two orbitals are still permitted, i.e., the local orbital quantum number is not conserved in the lattice. For two degenerate orbitals, all local Coulomb interactions can be expressed in terms of the Hubbard interaction U and the Hund's-rule coupling J. The Hubbard interaction suppresses charge fluctuations in the lattice, and the Hund's rule coupling tends to maximize the local spin. We incorporated the symmetry constraints in the Gutzwiller variational states.

As our first application, we studied the ferromagnetic phase transition as a function of the model parameters for various band fillings. In general, a large density of states and a strong Hund's-rule exchange favor ferromagnetism. In the Gutzwiller wave function, the ferromagnetic order is strongly suppressed so that much larger interaction strength are needed than predicted by the Hartree-Fock solution. Moreover, the regions in parameter space where non-saturated ferromagnetism occurs are much broader in Gutzwiller theory. As shown in earlier studies, this gives room for the experimental observations of non-saturated ferromagnetism, e.g., in transition metals such as nickel and iron. We find that, away from half band-filling, the ground state energy converges very quickly so that the first-order and second-order corrections to the results in infinite dimensions are small compared to the ground-state energy when the potential energy is of the same order as the kinetic energy. In general, the magnetization is only marginally affected by the higher orders so that the infinite-d approximations gives already reliable results.

As a second application, we investigated the interaction-induced deformation of the Fermi surface. These effects occur for large interaction strength, when the potential energy of the system is twice as large as the kinetic energy. For weaker interactions and small densities, the deformations of the Fermi surface can be neglected. Close to half band-filling and for special choices of the electron transfer parameters, the interactions can induce a change in the Fermi-surface topology from open to closed constant-energy contours. These effects are a result of the finite-order diagrams and cannot be seen in the Gutzwiller approximation because the orbital densities are fixed due to the lattice symmetries. The contributions beyond our second-order approximation still affect the Fermi surface and the density matrix. However, we can assume that the qualitative findings are valid in all orders of the approximation. To test this hypothesis, we analyzed a model of uncoupled bands and compared our results to a higher-order study. Furthermore, this model shows that systems with a tendency to noticeable Fermi-surface deformations can also break the lattice symmetry (Pomeranchuk effect). The energy gain in Gutzwiller theory due to Fermi-surface deformations is of the order of thermal energies at room temperature so that

the Pomeranchuk effect can be responsible for changes in the lattice structure.

4.2 Conclusions and outlook

In this work, we developed a systematic diagrammatic approach for the evaluation of expectation values for Gutzwiller-correlated multi-band wavefunctions. In principle, the inclusion of more complicated interactions such as nonlocal density-density interactions is straightforward. All necessary diagrams and external weights can be computed with the algorithms that are described in this work. However, the calculation of the second-order corrections is numerically demanding even for two orbitals. Employing symmetries made it possible to calculate all diagrams to second order. Since the number of diagrams rises exponentially as a function of the order and as a function of the number of orbitals, it is important to develop new strategies for a compact evaluation of finite orders, to study systems with more bands, and to investigate symmetry-broken states such as antiferromagnetism or superconductivity.

The analysis of our general diagrammatic approach showed that vertices with six or more nodes give a negligible contribution to a given order. Pictorially, a vertex with six lines corresponds to a three-particle scattering event for which the phase space is fairly small. Therefore, one may focus on diagrams with four-vertices (two-particle scatterings) only. Unfortunately, the exponential increase in diagrams makes it very difficult to go beyond third order in the expansion, even for a two-band model. Apparently, it is necessary to reduce the number of independent diagrams in the expansion. The large number of independent diagrams is primarily caused by the orbital mixing, i.e., at every point in the Brillouin zone the local orbitals mix differently to form the Bloch bands. Obviously, one strategy out of this dilemma is to study multi-band models where this mixing is absent by construction. Apparently, such independent-band models do not reflect the physics of transition metals and their compounds but they might be realizable in cold-atom setups.

The strong deformations of the Fermi surface show that a theory that incorporates the Fermi surface of the noninteracting system will fail for strong interactions and densities close to half filling. On the other hand, the optimization of the Fermi surface leads to an effective single-particle Hamilton operator that can be used as the starting point for Fermi liquid theory. For example, the stability of the magnetic ordering for finite temperatures could be analyzed within Landau-Gutzwiller theory. Moreover, the effective single-band Hubbard model shows, that strong Fermi surface deformations can lead to symmetry broken phases. In our two-band model, a broken spatial symmetry leads to different orbital densities which are energetically unfavorable. The kinetic and the potential energy increase for large values of J/U. It is an open question if an increase in energy can be compensated by an optimization of the Fermi surface.

In this work, we used a fixed model Hamiltonian and performed a study of the fundamental characteristics of our model. Alternatively, the infinite-d approximation can be used in a combination with Density Functional Theory to construct

an ab-initio calculation scheme. Our findings show that the infinite-d limit can be sufficient if the correlation effects lead to an adjustment of the orbital densities as in the ferromagnetic case.

Appendix

A.1 Proof of the consistency of the HF-mapping

A.1.1 Laplace formula

The Laplace formula for the expansion of the determinant $\det[A]$ of a matrix A can be written down in our index set notation. We can choose an arbitrary subset of columns $J_2 \subset I_2$ with $|J_2| = k$ and sum over all rows $J_1 \subset I_1$ with $|J_1| = k$

$$\det[A_{I_1,I_2}] = \sum_{J_1 \subset I_1, |J_1|=k} [\overleftarrow{J_1, I_1}]^+_\Sigma [\overleftarrow{J_2, I_2}]^+_\Sigma \det[A_{I_1 \setminus J_1, I_2 \setminus J_2}] \det[A_{J_1,J_2}] \; . \tag{A.1.1}$$

It is also possible to sum over the columns instead of the rows.

A.1.2 Derivation

The mapping from a HF-operator to normal ladder operators in section 2.2.1 should be the inversion of the expansion of an operator in terms of HF-operators in section 2.2.2. We prove this as follows.

We start by writing down the expansion of an operator in HF-operators as in Eq. (2.2.14),

$$\hat{C}^\dagger_{K_1} \hat{C}_{K_2} = \sum_{\substack{I_1 \subset K_1 \\ I_2 \subset K_2}} X^{K_1,K_2}_{I_1,I_2} \left(\hat{C}^\dagger_{I_1} \hat{C}_{I_2} \right)^{\mathrm{HF}} \tag{A.1.2}$$

with

$$X^{K_1,K_2}_{I_1,I_2} = [\overleftarrow{I_1, K_1}]^+_\Sigma [\overrightarrow{K_2, I_2}]^-_\Sigma \{ \hat{C}^\dagger_{K_1 \setminus I_1} \hat{C}_{K_2 \setminus I_2} \} \; , \tag{A.1.3}$$

and we insert the definition of the HF-operators which is given in the second line in Eq. (2.2.11). We find

$$\hat{C}^\dagger_{K_1} \hat{C}_{K_2} = \sum_{\substack{I_1 \subset K_1 \\ I_2 \subset K_2}} X^{K_1,K_2}_{I_1,I_2} \sum_{\substack{J_1 \subset I_1, J_2 \subset I_1 \\ |J_1|=|J_2|}} (-1)^{|J_1|} [\overrightarrow{I_1, J_1}]^+_\Sigma [\overleftarrow{J_2, I_2}]^-_\Sigma \{ \hat{C}^\dagger_{I_1} \hat{C}_{I_2} \} \hat{C}^\dagger_{I_1 \setminus J_1} \hat{C}_{I_2 \setminus J_1} \tag{A.1.4}$$

$$= \sum_{\substack{L_1 \subset K_1 \\ L_2 \subset K_2}} {\sum_{\substack{M_1 \subset K_1 \setminus L_1 \\ M_2 \subset K_2 \setminus L_2}}}' X^{K_1,K_2}_{L_1 \cup M_1, L_2 \cup M_2} (-1)^{|M_1|} [\widetilde{L_1, M_1}]^+_\Sigma [\widetilde{M_2, L_2}]^-_\Sigma \{ \hat{C}^\dagger_{M_1} \hat{C}_{M_2} \} \hat{C}^\dagger_{L_1} \hat{C}_{L_2} \; .$$

This can be summarized as

$$\hat{C}^\dagger_{K_1}\hat{C}_{K_2} = \sum_{\substack{L_1\subset K_1\\ L_2\subset K_2}} \hat{C}^\dagger_{L_1}\hat{C}_{L_2} \sum_{\substack{M_1\subset K_1\setminus L_1\\ M_2\subset K_2\setminus L_2}}{}' (-1)^{|M_1|} S^{L_2,K_2,M_2}_{L_1,K_1,M_1}$$

with

$$\begin{aligned} S^{L_2,K_2,M_2}_{L_1,K_1,M_1} =& \{\hat{C}^\dagger_{M_1}\hat{C}_{M_2}\}\{\hat{C}^\dagger_{K_1\setminus(L_1\cup M_1)}\hat{C}_{K_2\setminus(L_2\cup M_2)}\} \\ & \widetilde{[L_1,M_1]}^+_\Sigma[\widetilde{M_2,L_2}]^-_\Sigma[\overleftarrow{L_1\cup M_1,K_1}]^+_\Sigma[\overrightarrow{K_2,L_2\cup M_2}]^-_\Sigma \,. \end{aligned} \tag{A.1.5}$$

It is easy to see that

$$\begin{aligned} \widetilde{[L_1,M_1]}^+_\Sigma[\overleftarrow{L_1\cup M_1,K_1}]^+_\Sigma &= [\overleftarrow{M_1,M_1\cup L_1}]^+_\Sigma[\overleftarrow{L_1\cup M_1,K_1}]^+_\Sigma \\ &= [\overleftarrow{M_1,K_1\setminus L_1}]^+_\Sigma[\overleftarrow{L_1,K_1}]^+_\Sigma \,, \end{aligned} \tag{A.1.6}$$

so that we can rewrite Eq. (A.1.5) as

$$S^{L_2,K_2,M_2}_{L_1,K_1,M_1} = [\overleftarrow{L_1,K_1}]^+_\Sigma[\overrightarrow{K_2,L_2}]^-_\Sigma \det[\rho_{M_1,M_2}]\det[\rho_{(K_1\setminus L_1)\setminus M_1,(K_2\setminus L_2)\setminus M_2}] \,, \tag{A.1.7}$$

where we switched to the determinant representation of the internal contractions. Now, we can employ Laplace's formula Eq. (A.1.1). We sum over all subsets M_2 with $|M_2| = |M_1| = k =$ fixed and find

$$\hat{C}^\dagger_{K_1}\hat{C}_{K_2} = \sum_{\substack{L_1\subset K_1\\ L_2\subset K_2}}{}' \hat{C}^\dagger_{L_1}\hat{C}_{L_2}[\overleftarrow{L_1,K_1}]^+_\Sigma[\overrightarrow{K_2,L_2}]^-_\Sigma \sum_{M_1\subset K_1\setminus L_1} (-1)^{|M_1|} S^{L_2,K_2}_{L_1,K_1;|M_1|}$$

$$S^{L_2,K_2}_{L_1,K_1;k} := \sum_{\substack{M_2\subset K_2\setminus L_2\\ |M_2|=k}} \det[\rho_{M_1,M_2}]\det[\rho_{(K_1\setminus L_1)\setminus M_1,(K_2\setminus L_2)\setminus M_2}] \tag{A.1.8}$$

$$= [\overleftarrow{L_1,K_1}]^+_\Sigma[\overrightarrow{K_2,L_2}]^-_\Sigma \det[\rho_{K_1\setminus L_1,K_2\setminus L_2}] \,. \tag{A.1.9}$$

The value of $S^{L_2,K_2}_{L_1,K_1;k}$ does not depend on k so that we have an alternating sum over all subsets $|M_1|$

$$\hat{C}^\dagger_{K_1}\hat{C}_{K_2} = \sum_{\substack{L_1\subset K_1\\ L_2\subset K_2}}{}' \hat{C}^\dagger_{L_1}\hat{C}_{L_2}[\overleftarrow{L_1,K_1}]^+_\Sigma[\overrightarrow{K_2,L_2}]^-_\Sigma\{\hat{C}^\dagger_{K_1\setminus L_1}\hat{C}_{K_2\setminus L_2}\} \sum_{M_1\subset K_1\setminus L_1} (-1)^{|M_1|} \tag{A.1.10}$$

Then, we can use

$$\sum_{J\subset K} (-1)^{|J|} = \begin{cases} 1 & K = \varnothing \\ 0 & \text{otherwise} \end{cases} \tag{A.1.11}$$

which follows from ($n \geqslant 1$)

$$\sum_{k=0}^{n}(-1)^k\binom{n}{k} = (1-1)^n = 0\,. \tag{A.1.12}$$

Eq. (A.1.10) and the constraint $|K_1\backslash L_1| = |K_2\backslash L_2|$ assert that the only nonvanishing contribution will come from the case $K_1 = L_1$ and $K_2 = L_2$. Thus, we find that both sides from Eq. (A.1.10) are equivalent,

$$\hat{C}^\dagger_{K_1}\hat{C}_{K_2} = \sum_{\substack{L_1\subset K_1\\ L_2\subset K_2}} \hat{C}^\dagger_{L_1}\hat{C}_{L_2}\delta_{L_1,K_1}\delta_{L_2,K_2}\,, \tag{A.1.13}$$

by using $\{\hat{C}^\dagger_\varnothing\hat{C}_\varnothing\} = 1$ by definition.

A.2 Program details

A.2.1 Generation of diagrams

The generation of diagrams is carried out in four steps. As a first step, we generate all connected diagrams. At this stage, a diagram will be defined by a set of internal nodes $z_1, \ldots, z_n$ which must be connected to one or two external nodes m_1 and m_2. The nodes are either connected or not connected, so that the 'connectivity mask' of these nodes can be represented by a single binary number. The node is placed in the order $m_1, m_2, z_1, z_2, \ldots$ if two external nodes are present and in the order $m_1, z_1, z_2, \ldots$ if only one external node is present. The first bit of the binary number represents the connectivity between node 1 and 2. The next two bits represent the connectivity between node 1 and 3 and between nodes 2 and 3 and so on. We collect all connectivity masks which are mapped to each other by a permutation of the internal nodes. An example is given in Fig. A.1.

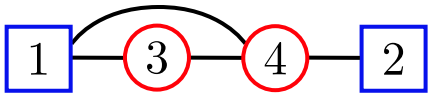


Figure A.1: Connected diagram. The connectivity mask of this diagram is 111010 where the first bit is the rightmost bit.

In the second step, we replace the symbolic nodes by all possible external nodes $Z_{I_1,I_2}\tilde{C}^\dagger_{I_1}\tilde{C}_{I_2}$ and all internal nodes $M_{I_1,I_2}\tilde{C}^\dagger_{I_1}\tilde{C}_{I_2}$. The number of external nodes can be reduced by the incorporation of the symmetries of the system. After we have fixed the external nodes, we start to replace the connectivity lines by all possible multi-lines

$$\mathrm{ml}_n = \mathrm{ml}_n(I_1, I_2, J_1, J_2) = \{\tilde{C}^\dagger_{i,I_1}\tilde{C}_{i,I_2}\tilde{C}^\dagger_{j,J_1}\tilde{C}_{j,J_2}\}\,, \tag{A.2.1}$$

where n gives the index of the line defined by the nth nonzero bit (from right to left). The multi-lines define which operators of the nodes on site i and j are contracted

with each other. In order to avoid the explicit introduction of the index n, we can save the multi-lines in an array, and use the same ordering as for the connectivity lines. Therefore, the multi-line depends only on the index set of the operators at this stage. The multi-lines define all Graßmann variables of the internal nodes. In this way, the iteration over all multi-lines automatically generate the internal nodes. A diagram that contains an internal node which is not allowed is sorted out immediately. This algorithm offers a fast and simple way to generate all possible diagrams in the case of a nondiagonal density matrix $\tilde{\rho}_{ij}^{\sigma\sigma'}$.

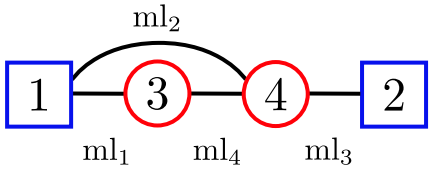


Figure A.2: Connected diagram with multi-lines.

In a third step, we split up the multi-lines and use the symmetries of the density matrix and of the coefficients Z_{I_1,I_2}. Therefore, we have to reorder all Graßmann variables which define our diagram. We shift the operators in a way so that the Graßmann variables which form a multi-line stand next to each other and evaluate the overall change in the sign of our diagram. Then, we evaluate the contractions symbolically so that a multi-line represents a product of entries of the density matrix rather than a subdeterminant. In the spin-band symmetric case for example, each multi-line can be written as a function of the three exponents

$$\mathrm{ml}_n = \mathrm{ml}(m_1, m_2, m_3) = \left(\tilde{\rho}_{ij}^{11}\right)^{m_1} \left(\tilde{\rho}_{ij}^{12}\right)^{m_2} \left(\tilde{\rho}_{ij}^{22}\right)^{m_2} \quad . \tag{A.2.2}$$

At this point we can also combine equal coefficients Z_{I_1,I_2} and summarize all equal diagrams. Note that we must also keep the index of the external and internal nodes after we used the symmetry relations. After this step, we save all topological different diagrams to disk before we proceed any further.

In a fourth and last step, we align the diagrams on a square lattice. The maximal distance between two nodes is defined by the cutoff length dmax so that we drop all entries of the density matrix $\tilde{\rho}_{ij}$ with $||i-j||_1 \leqslant \mathrm{dmax}$, where $||.||_1$ gives the one norm (Manhatten metric). This restriction does not account for the multiplicity of the lines between the nodes but the set of possible alignments depends on the connectivity mask only. Therefore, this computational step has to be done only once for a large number of diagrams. For each connectivity mask, we generate all possible alignments. Then, we combine all geometrical variations of the diagram which can be obtained by the action of an elements of the D_4 lattice symmetry (all possible rotations and reflections). In order to do so, we map all displacement vectors r_{ij} into the first octant such that $0 \leqslant r_y \leqslant r_x$, and save the symmetry transformation which is used. Then, we combine the lattice alignments with our diagrams. The multi-line is now combined with the index i_r of the distance vector r_{ij}

$$\mathrm{ml} = \mathrm{ml}(i_r, m_1, m_2, m_3) = \left(\tilde{\rho}_{ij}^{11}\right)^{m_1} \left(\tilde{\rho}_{ij}^{12}\right)^{m_2} \left(\tilde{\rho}_{ij}^{22}\right)^{m_2} \quad , \tag{A.2.3}$$

where the subindex of the multi-line can be dropped. This means that the ordering of the multi-lines does not matter anymore. At this point, we have to use the symmetry transformation and apply it to the density-matrix elements. Additionally, we can use the multiplicity of the lines between two nodes to restrict the diagrams. For example, we can restrict the total length of the diagrams by

$$\mathrm{dmax}' = \sum_k w(r_k)^{m_1+m_2+m_3} \ , \tag{A.2.4}$$

where the sum over k runs over all multi-lines and the weight $w(r_k)$ defines the importance of the density matrix entries with distance vector r_k. One can easily think of several possibilities to vary these restrictions. However, the number of diagrams will increase significantly if the lattice restrictions are relaxed in any way. Therefore, we use a simple cutoff length with dmax = 4 or dmax = 5 in most cases. Now, it does not matter anymore to which node a multi-line belongs to so that we can summarize all equal diagrams again. For example, it is possible that two nodes have the same distance vector r_i so that two different diagrams result in the same monomial of the Z_{I_1,I_2} and the $\rho_{ij}^{\sigma\sigma'}$ coefficients. All diagrams are saved to disk again.

A second-order expansion of the band symmetric setup leads to 4 Gigabyte of data for dmax = 4 and 20 Gigabyte of data for dmax = 5. These numbers are the upper limit for a practical evaluation of the diagrams with respect to the available resources. The number of diagrams needed for the computation of the kinetic energy in nth order is usually as large as for the potential energy in order $n+1$. The number of diagrams will increase drastically if the lattice and orbital symmetries are dropped or if we increase the number of bands or the number of internal nodes. In order to keep the number of diagrams as small as possible, we constructed the algorithm in such a way that symmetries of the system are built in as early as possible. In principle, the algorithm provides a way to construct all diagrams for an arbitrary number of internal nodes and an arbitrary number of bands. The lattice geometry can be adapted easily.

A.2.2 Evaluation of diagrams

The evaluation of the diagrams consists of two main steps. First, we have to read in all diagrams for a given density matrix ρ and replace the lines with their numeric value. In the generation of the diagrams we reduced the external nodes by an application of the symmetries of the system. Now, we can construct the remaining nodes by the action of the inverse symmetry operation. The symmetry operation must be applied to the lines of the diagram as well as to its internal nodes. Thereby, the sign of the diagram may change. The reduction of the internal nodes in the generation of the diagrams is consistent with the action of the symmetry operation. This means that the symmetry operations must lead to the same change in the sign for all internal nodes which were summarized.

After this step, we have produced a polynomial that depends on the internal and external nodes. The advantage of this procedure is that we can now use the polynomial to perform an optimization of the variational parameters in a second step.

This reduces the computational cost of the optimization tremendously. However, the evaluation of the gradient with respect the line $\rho_{ij}^{\sigma\sigma'}$ still requires the evaluation of all diagrams.

A.2.3 Optimization of the variational parameter

The optimization of the variational parameter in the presence of the gauge constraints in Eq. (2.4) can be carried out in several ways. Here, we list only three of them which can be combined to achieve optimal results. In all cases, the symmetries of the system should be used to reduce the number of free parameter. The first approach is very similar to the optimization procedure defined in [28]. If we do not have made any other guess for the parameter set, we always start our optimization in $\lambda =$ Id. For a given density matrix ρ, we can evaluate the gradient of the energy $E_G = \langle \hat{H} \rangle_G$ with respect to the λ_{I_1,I_2} parameters analytically. Then, the gradient is projected to the tangential space $T_\lambda M$ of the submanifold M which is defined by the gauge constraints. We set

$$\widetilde{\nabla_\lambda E_G}\Big|_{\lambda^n} = \mathrm{Pr}_{\lambda^n}(\nabla_\lambda E_G) \ , \tag{A.2.5}$$

with

$$\mathrm{Pr}_\lambda(v+w) = v \ , \quad v \in T_\lambda M, \ w \in T_\lambda^\perp M \ , \tag{A.2.6}$$

where $T_\lambda^\perp M$ is the orthogonal complement of the tangential space. In order to compute the projection numerically, we have to build up an orthonormal basis $\{v_i^\perp\} \in T_\lambda^\perp M$ from the basis vectors which are given by the gradients of the constraint function $\phi_i(\lambda) = 0 \ , \ \forall \lambda \in M$. We set

$$\phi_0 = Z_\varnothing - 1 \ , \tag{A.2.7}$$

$$\phi_i = Z_{\sigma_i,\sigma_i'} \ , \tag{A.2.8}$$

where we used the normal ordering as index function $i \to (\sigma_i, \sigma_i')$. Then, we make a step along the direction of the gradient

$$\lambda^{(n+1)\prime} = \lambda^{(n)} + \alpha_n \mathrm{Pr}_{\lambda^{(n)}}(\nabla_\lambda E_G) \ , \tag{A.2.9}$$

with a suitable choice of α so that the deviation to the submanifold is not too large. After that, we have to project the new point $\lambda^{(n+1)\prime} \to \lambda^{(n+1)}$ down to the submanifold M. Several strategies can be used to do so. The simplest one is to apply a numerical minimization of the penalty function $h = \sum_i \phi_i^2$. The projection will be well defined as long as the deviations from the submanifold are not large. Alternatively, we can use the orthonormal vectors in $\lambda^{(n)}$ to restrict the 'direction' of the projection. Therefore, we have to solve the following problem

$$h(\lambda^{(n+1)\prime} + \sum_i v_i^\perp \mu_i) = 0 \ . \tag{A.2.10}$$

The coefficients can obtained by a numerical minimization of the coefficients μ_i. This problem is solved approximately in [28], but we can also solve this problem with high

numerical precision with low computational costs. However, this procedure does not seem to have any benefits over a direct minimization of the penalty function. We implement the minimization with the help of the Gnu Scientific Library [37] which offers several numerical routines to choose the parameter α appropriately. In principle, this minimization scheme can be combined with any optimization algorithm which depends on the gradient of the function only.

In the second approach, we solve the gauge constraints exactly by a restriction of all λ_{I_1,I_2} with $|I_1| = |I_2| < 2$. We use that the single-particle sector $\tilde{\lambda}^1$ of the variational parameters is diagonal so that $\lambda_{\sigma\sigma'} \sim \delta_{\sigma\sigma'}$ and $\tilde{\lambda}_\sigma = \lambda^2_\sigma$. The Z_{I_1,I_2} coefficients can always be written as a linear function of the $\tilde{\lambda}_{I_1,I_2}$ coefficients. We can express the constraint functions as

$$\phi(\lambda_{I_1,I_2}) = A\tilde{\lambda}^1 + B\lambda^r \;, \tag{A.2.11}$$

where λ^r includes all terms with $|I_1| = |I_2| > 1$ and $\tilde{\lambda}^1$ shall also include $\tilde{\lambda}_\varnothing$. Then, we have to solve the linear problem

$$A\tilde{\lambda}^1 = C \;, \tag{A.2.12}$$

where we set $C = -B\lambda^r$. The matrix A depends only on the local density matrix. We can also use this relation to obtain the derivatives of $\tilde{\lambda}^1$ with respect to the free variational parameters $\partial_\nu \lambda^1_\mu = \partial\lambda^1_\mu/\partial\lambda^r_\nu$. Therefore, we take the derivative of Eq. (A.2.12) and solve the equations

$$A\,\left(\partial_\nu \tilde{\lambda}^1\right) = \partial_\nu C \tag{A.2.13}$$

and set $\partial_\nu \lambda^1_\mu = (\partial_\nu \tilde{\lambda}^1)/(2\lambda^1_\mu)$. Of course, we will have to set the gradient to zero if the corresponding parameter is set to zero by hand in order to incorporate the symmetries and conservation rules of the system. Therefore, we define the set I_0 of all nonzero variational parameters and set $\partial_\nu \tilde{\lambda}^1_\mu = 0$ for all $\mu \in I_0$. The main drawback of this procedure is that the free parameters still have to obey the inequality constraints $\tilde{\lambda} > 0$ after we solved Eq. (A.2.12). It is possible that the optimization algorithm leaves the valid parameter region. In most cases, this difficulty can be overcome if we apply the first optimization algorithm to obtain an approximate solution and refine our result with the second optimization algorithm. The exact evaluation of the constraint function can be important for the precision of the Gutzwiller energy E_G.

A third option is to solve the equality constraints exactly and employ an penalty function which is as defined as

$$g(\lambda) = \begin{cases} 0 & \text{if } (\tilde{\lambda}^1)_\mu > 0 \;, \quad \forall\, \mu \in I_0 \\ g_0 & \text{otherwise} \end{cases} \tag{A.2.14}$$

with some large value of g_0. This optimization algorithm is the most robust algorithm but it converges only slowly.

A.2.4 Important diagrams

In Fig. A.3, we show the contributions of different diagrams to the Gutzwiller energy. The energy is evaluated for the optimized variational parameters for the parameters $t_x^{11} = -1.0$, $t_y^{11} = -0.6$, $t_{xy}^{11} = 0.2$, $t_{xy}^{12} = -0.4$, $U = 6.0$, $J = 1.6$ and $n_0 = 0.4$. The diagrams are grouped by the number of lines attached to their internal and external nodes. The most important contributions come from the infinite-d terms. However, the optimization of the Fermi surface depends strongly on the nontrivial diagrams which contain the long-ranged lines.

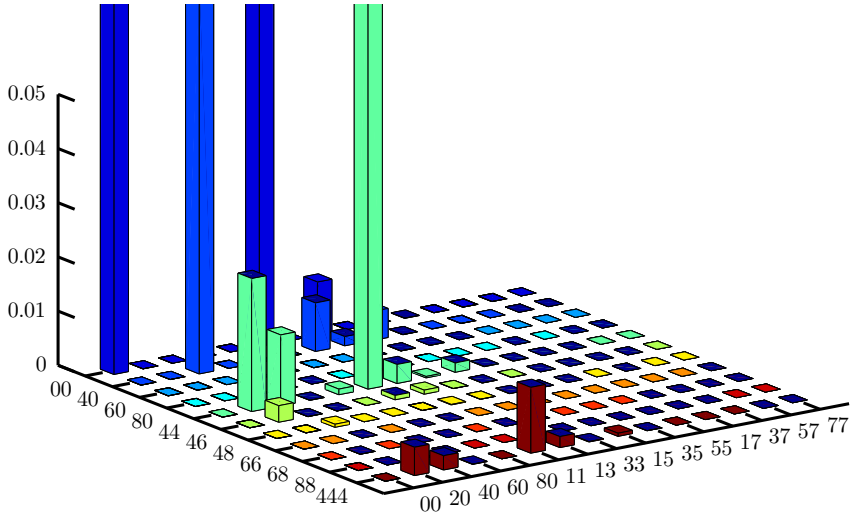


Figure A.3: Contributions of diagrams to the Gutzwiller energy E_G. The diagrams are sorted by the number of lines of the external and internal nodes. Negative values are marked with a dark blue top area. The model parameters are set to $t_x^{11} = -1.0$, $t_y^{11} = -0.6$, $t_{xy}^{11} = 0.2$, $t_{xy}^{12} = -0.4$, $U = 6.0$ and $J = 1.6$. The variational parameters have been optimized for $n_0 = 0.4$. The most important non-trivial contributions are given by diagrams with internal nodes with only four lines.

All contributions with $|E_G| > 10^{-3}$ are shown in Fig. A.4. The total number of all lines in a diagram is given by the sum of all lines attached to the nodes divided by two. This shows that the weight of all combined diagrams with a specific number of lines does not decrease monotonically. All contributions have internal nodes with four internal lines except for the diagrams with internal nodes with 4 and 6 lines and an external node with two lines. Therefore, it is reasonable to argue that only the diagrams with the internal nodes Z_{I_1,I_2} with $|I_1| = 2 = |I_2|$ need to be considered in the expansion. We included these third-order terms in the calculations of the energy shown above. However, these terms generate 62 Gigabyte of data for dmax = 4 for the two-band system with spin-band symmetry. Therefore, a repeated evaluation of all diagrams will cost too much time. Additionally, the analysis of higher orders of the single-band case discussed in section 3.3.4 shows that the first and the third order usually give unreasonable results that deviate strongly from the infinite-d approximation and the corresponding even orders. Therefore, we expect that even a fourth-order expansion would be necessary to increase the accuracy of our results.

Therefore, we would have to implement a further restriction of the external nodes to reduce the number of lines.

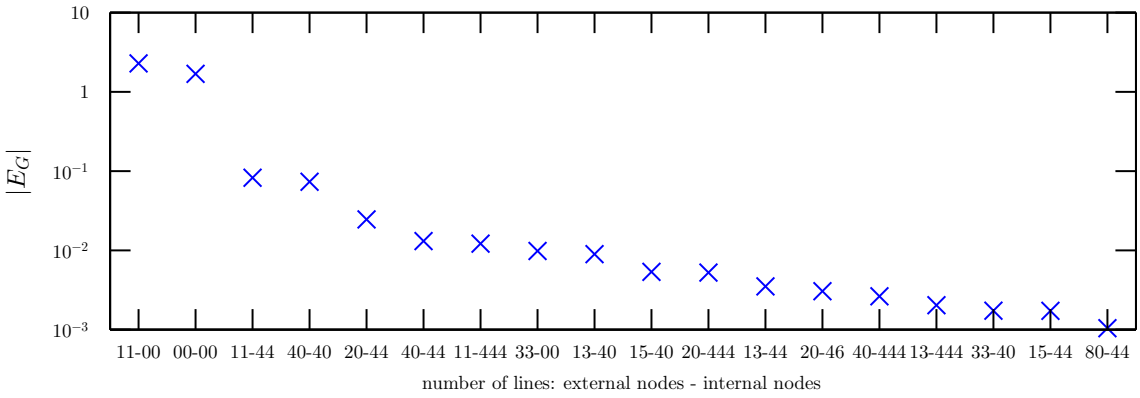


Figure A.4: The absolute value of the Gutzwiller energy $|E_G|$ is given for the most important contributions of the diagrammatic expansion. The diagrams are sorted by the number of lines of the external and internal nodes. The model parameters are set to $t_x^{11} = -1.0$, $t_y^{11} = -0.6$, $t_{xy}^{11} = 0.2$, $t_{xy}^{12} = -0.4$, $U = 6.0$ and $J = 1.6$. The variational parameters have been optimized for $n_0 = 0.4$.

In Fig. A.5, we show the number of diagrams for the different groups of internal and external nodes. Even though the number of diagrams increases dramatically if the external and internal nodes have a large number of lines, their total weight will be negligible. The multiplication of all lines in a diagram reduces its weight strongly. An analysis of the energy shows that the contributions from diagrams with a very large number of lines will remain small even if the weights of the external and internal nodes are not taken into account. Nevertheless, we can neglect all diagrams with very small external weights because the line weight already reduces the weight of all diagrams to a finite value. For example, we can neglect the evaluation of the diagrams with an external weight $w_{\text{ext}} < 10^{-10}$. The internal nodes do not necessarily reduce the weight of a diagram. Usually, some of the Z_{I_1,I_2} coefficients are between $1 < Z_{I_1,I_2} < 2$ for $|I_1| = 2 = |I_2|$. The external nodes with six lines are quite small so that $Z_{I_1,I_2} \approx 0.1$ for $|I_1| = 3 = |I_2|$ but the external node with eight lines can be close to unity. For example, we get $Z_{u_1 d_1 u_2 d_2} = -1.0248$ after the optimization of the system that we discussed above.

The number of all relevant diagrams for a given contribution to $|E_G|$ is listed in the following table (dmax = 4). We summarize all diagrams in the categories shown in Fig. A.3 and neglect those which have an contributions which lies below the threshold $|E_G^{\text{cat.}}|$.

$\lvert E_G^{\text{cat.}}\rvert \geqslant$	2nd order	3rd order & $\lvert I_{1,2}\rvert = 2$
10^{-2}	$4.5\ 10^4$	$2.8\ 10^6$
10^{-3}	$1.4\ 10^6$	$9.8\ 10^7$
10^{-4}	$5.2\ 10^6$	$3.2\ 10^8$
10^{-5}	$3.2\ 10^7$	$6.8\ 10^9$
10^{-6}	$8.7\ 10^7$	$1.3\ 10^{10}$
0	$5.2\ 10^8$	$1.7\ 10^{10}$

Table A.1: Number of diagrams in the categories shown in Fig. A.3. We neglect all categories which have an contributions which lies below the threshold $|E_G^{\text{cat.}}|$.

The second-order results contain all diagrams in second order while the third-order results also contain all third order diagrams with internal nodes that have four lines.

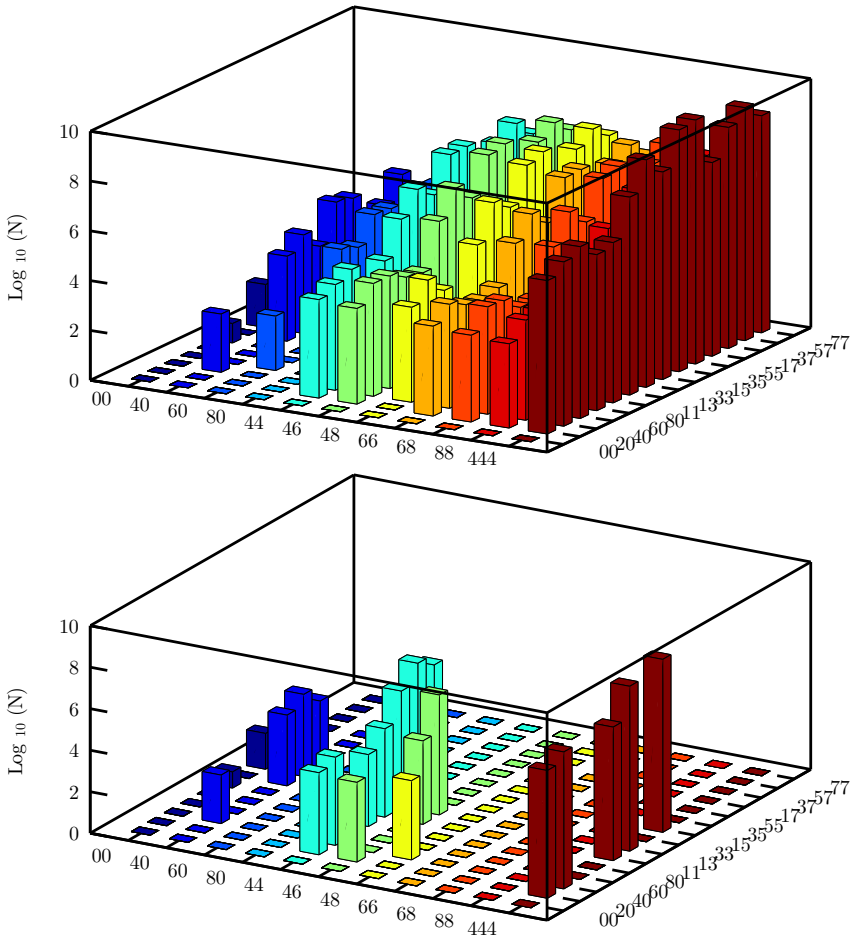


Figure A.5: Number of diagrams for the spin band symmetric p_x-p_y two band model. The diagrams are grouped by the number of lines of their internal and external nodes. The cutoff parameter of the density matrix is set to dmax = 4. The upper panel shows the number of all categories. The lower panel shows only the number of diagrams for the categories that have an contribution $|E_G^{\text{cat.}}| \geqslant 10^{-4}$ to the energy.

A.3 Hamilton operator

A.3.1 Crystal field

The electronic orbitals of an atom placed in a crystal lattice are affected by the Coulomb interaction with the nearest neighbors which are called 'ligands'. In general, the atomic orbitals are no longer eigenstates of the Hamilton operator that includes the field of the neighboring atoms. In crystal-field theory, the ligands are considered to be monopols that attract the electrons on its original site. In this way we can determine a set of new eigenstates that can be expressed as a simple superposition of the original orbitals. The following derivations can be found in [27] but are repeated here to illustrate the concept. Additionally, a few results are added to cover the case of a square lattice.

We start by writing down the definition of the atomic orbitals

$$\phi_{n,l,m}(r,\theta,\varphi) = R_{nl}(r)Y_{lm}(\theta,\varphi) \ , \tag{A.3.1}$$

where n is the principal quantum number and l the absolute value of the angular momentum and m its z-component. The wave function can be split up into a radial part $R_{nl}(r)$ and a angular dependent part $Y_{lm}(\theta,\varphi)$ which are defined in terms of Laguerre and Legendre polynomials, respectively. In order to evaluate the potential energy of the electrons we need to transform the distance vector r_{ei} between the electron and the neighboring atoms to the coordinate basis of the initial atom. In order to do so, we need the addition theorem of spherical harmonics which is given without proof [38]. Let ζ and η be two points on the unit sphere S^2, then

$$\sum_{l,m} Y_{lm}(\zeta)Y^*_{lm}(\eta) = \frac{2m+1}{4\pi}P(\cos(\theta_{12})) \ , \tag{A.3.2}$$

where $P(\cos(\theta_{12}))$ gives the corresponding Legendre polynomial in terms of the angle θ_{12} between the vectors η and ζ. A Taylor expansion of the inverse of the distance r_{12} between the vector r_1 and r_2 gives

$$\frac{1}{r_{12}} = \sum_{l=0}^{\infty} \frac{\min(r_1,r_2)^l}{\max(r_1,r_2)^{l+1}} P_l(\cos(\theta_{12})) \ . \tag{A.3.3}$$

Now we express the distance vector r_{12} as

$$\frac{1}{r_{12}} = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} \frac{\min(r_1,r_2)^l}{\max(r_1,r_2)^{l+1}} \left(\frac{4\pi}{2l+1}\right)^{1/2} Y^*_{lm}(\theta_1,\varphi_1) \left(\frac{4\pi}{2l+1}\right)^{1/2} Y_{lm}(\theta_2,\varphi_2) \tag{A.3.4}$$

$$= \sum_{l=0}^{\infty} f_l(r_1,r_2) \sum_{m=-l}^{l} (-1)^m Y_{lm}(\theta_1,\varphi_1)Y_{lm}(\theta_2,\varphi_2) \ .$$

When the lattice spacing a is much larger than the length scale of the atomic orbitals, we can safely assume that the distance r_e to the electron is smaller than the distance

to the ligands $r_i = a$. Thus, we can make the approximation

$$\frac{\min(r_e, r_i)^l}{\max(r_e, r_i)^{l+1}} \approx \frac{r_e^l}{a^{l+1}} \; . \tag{A.3.5}$$

The potential energy $V(\mathbf{r})$ of the electron in the ligand field can be written as

$$V(\mathbf{r}) = \sum_i \frac{1}{r_{ei}} = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} r^l Q_{l,m}(a) \left(\frac{4\pi}{2l+1} \right)^{1/2} Y_{lm}(\theta, \varphi) \; . \tag{A.3.6}$$

The coefficients

$$Q_{l,m}(a) = \left(\frac{4\pi}{2l+1} \right)^{1/2} \frac{1}{a^{l+1}} \sum_i Y_{lm}^*(\theta_i, \varphi_i) \; , \tag{A.3.7}$$

where $(\theta_i, \varphi_i) \in \{(\pi/2, 0), (\pi/2, \pi/2), (\pi/2, \pi), (\pi/2, 3/2\pi)\}$ give the angles of the neighboring atom positions. The charge of the electron and the dielectric constant have been set to unity. Now, the overlap integrals between two orbitals ϕ_1, ϕ_2 in the presence of the ligand field V can be evaluated as

$$\begin{aligned} V_{s_1 s_2} &= \int \mathrm{d}^3 r \phi_{s_1}^*(\mathbf{r}) V(\mathbf{r}) \phi_{s_2}(\mathbf{r}) \\ &= \sum_{l=0}^{\infty} \sum_{m=-l}^{l} Q_{l,m}(a) \left\langle n_{s_1}, l_{s_1} | r^{l+2} | l_{s_2}, n_{s_2} \right\rangle C^{l,m}(n_{s_1}, l_{s_1}; n_{s_2}, l_{s_2}) \; , \end{aligned} \tag{A.3.8}$$

with

$$C^{l,m}(m_{s_1}, l_{s_1}; m_{s_2}, l_{s_2}) = \left(\frac{4\pi}{2l+1} \right)^{1/2} \int Y(l,m) Y^*(l_{s_1}, m_{s_1}) Y(l_{s_2}, m_{s_2}) \;\; \mathrm{d}S^2 \; . \tag{A.3.9}$$

The conservation of angular momentum leads to the constraints that the C-coefficients are only nonvanishing if

$$\begin{aligned} & m = m_1 - m_2 \; , \\ & l + l_1 + l_2 = \text{even} \; , \\ & |l_1 - l_2| \leqslant l \leqslant l_1 + l_2 \; . \end{aligned} \tag{A.3.10}$$

These constraints show that only a finite number of terms in the summation over l is needed in the computation of the energy. The lattice structure is incorporated in the coefficients Q. The expectation values for the moments r^{l+2} can be evaluated analytically, as well as the coefficients C and Q.

We consider $3d$ orbitals on a square lattice with $n = 3$ and $l = 2$ and $m_1, m_2 \in \{-2, -1, 0, 1, 2\}$. The explicit calculation of the coefficients are carried out with a short Mathematica program [39]. The integrals over the radial part can be evaluated exactly

$$I_R(0) = 1, \; I_R(2) = 126, \; I_R(4) = 25515 \; . \tag{A.3.11}$$

The eigenvectors of the matrix V_{m_1,m_2} are

$$\begin{aligned} d_{x^2-y^2} &\sim \phi^{3d}_{-2} + \phi^{3d}_{-2}, \quad d_{z^2} \sim \phi^{3d}_{0}, \\ d_{xy} &\sim \phi^{3d}_{-2} - \phi^{3d}_{-2}, \quad d_{xz} \sim \phi^{3d}_{1} + \phi^{3d}_{-1}, \\ d_{yz} &\sim \phi^{3d}_{1} - \phi^{3d}_{-1} \, . \end{aligned} \tag{A.3.12}$$

The orbitals d_{xz} and d_{yz} are degenerate because they can be transformed into each other with a rotation of $\pi/2$ which leaves the lattice invariant.

In the next step we compute the differences between the eigenenergies and rescale these differences by the energy gap in units of $\Delta = E(d_{x^2-y^2}) - E(d_{z^2})$. We find

$$\begin{aligned} \Delta_{d_{xz},d_{yz}} &= 0 \, , \\ \Delta_{d_{xz},d_{z^2}} &= \frac{2025 - 4a^2}{12a^2 + 3375} \, , \\ \Delta_{d_{z^2},d_{xy}} &= \frac{4a^2 - 450}{4a^2 + 1125} \, , \\ \Delta_{d_{xy},d_{x^2-y^2}} &= 1 \, . \end{aligned} \tag{A.3.13}$$

It is easy to see that in the limit $a \to \infty$ the energy difference $\Delta_{d_{xz},d_{z^2}} < 0$, so that the d_{z^2} orbital becomes the energetically lowest orbital. For finite values of the lattice spacing, about $\approx 10 - 15$ times the Bohr radius a_{B}, the d_{xz} and d_{yz} values drop below the d_{z^2} orbital. Therefore, for realistic lattice spacings, the d_{xz} and d_{yz} orbitals are energetically below the d_{z^2} orbital. The qualitative behavior of these results coincide with the term schemes given in [40].

A.3.2 Interaction energy

The calculation of the on-site interaction energy in (3.1.13) requires the evaluation of integrals of the form

$$\begin{aligned} I^{s_2 s_3}_{s_1 s_4} &= \int\int \phi^*_{s_1}(r_a)\phi^*_{s_2}(r_b)\frac{1}{r_{ab}}\phi_{s_3}(r_b)\phi_{s_4}(r_a) \, \mathrm{d}V_a \, \mathrm{d}V_b \\ &= \int\int Y^*_{s_1}(\theta_a,\varphi_a)Y^*_{s_2}(\theta_b,\varphi_b)Y_{s_3}(\theta_b,\varphi_b)Y_{s_4}(\theta_a,\varphi_a)\frac{R^2(r_a)R^2(r_b)}{r_{ab}} \, \mathrm{d}V_a \, \mathrm{d}V_b \, . \end{aligned} \tag{A.3.14}$$

As described in the previous section, we need to express the distance vector r_{ab} in spherical harmonics. In a next step, we can express these integrals by the coefficients defined in Eq. (A.3.9) with the help of the relation

$$Y^*_{l,m} = (-1)^m Y_{l,-m} \, . \tag{A.3.15}$$

These integrals can be simplified to

$$I^{s_2 s_3}_{s_1 s_4} = \sum_{l=0}^{\infty} SC(l) \sum_{m=-l}^{l} C^{lm}(l_{s_1}, m_{s_1}, l_{s_4}, m_{s_4}) C^{l,-m}(l_{s2}, m_{s_2}, l_{s_3}, m_{s_3})(-1)^m \, , \tag{A.3.16}$$

where the Slater-Condon integrals

$$SC(l) = 4\pi \int \mathrm{d}r_a \int \mathrm{d}r_b \frac{\min(r_a, r_b)^l}{\max(r_a, r_b)^{l+1}} R^2(r_a) R^2(r_b) \tag{A.3.17}$$

are defined to shorten the notation. The coefficients in Eq. (A.3.16) are only nonvanishing if $m_{s_1} + m_{s_2} = m_{s_3} + m_{s_4}$ because of the constraints in Eq. (A.3.10). In order to construct the coefficients $(U_{\mathrm{int}})_{I_1,I_2}$, we have to express the p_x-p_y orbitals as linear combinations of the orbitals ϕ_{s_i}. This leads to a two-particle basis transformation

$$|J\rangle = \sum_I W_{J,I} |I\rangle \; . \tag{A.3.18}$$

Then we transform the coefficients $I^{s_2 s_3}_{s_1 s_4}$ with $J_1 = \{s_1, s_2\}$ and $J_2 = \{s_3, s_4\}$. We can cast the result into the form

$$U_{\mathrm{int}} = U \; \mathrm{Id} + \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & J \\ 0 & -3J & 0 & 0 & 0 & 0 \\ 0 & 0 & -2J & -J & 0 & 0 \\ 0 & 0 & -J & -2J & 0 & 0 \\ 0 & 0 & 0 & 0 & -3J & 0 \\ J & 0 & 0 & 0 & 0 & 0 \end{pmatrix} , \tag{A.3.19}$$

where we used the standard ordering of the two-particle sets defined in section 2.1. This result holds for p_x-p_y orbitals as well as for d_{xz}-d_{yz} orbitals. In general, any two orbitals which are the superposition of two atomic orbitals with the same n and l but opposite m will yield the same result due to the constraints in Eq. (A.3.10).

The specific values of U and J are given by

$$U = \frac{1}{49}(49SC(0) - 5SC(2)) \; , \tag{A.3.20}$$

$$J = -\frac{3}{49} SC(2) \; , \tag{A.3.21}$$

for p_x-p_y orbitals and

$$U = SC(0) + \frac{4SC(2)}{49} + \frac{4SC(4)}{49} \; , \tag{A.3.22}$$

$$J = \frac{3SC(2)}{49} + \frac{20SC(4)}{441} \tag{A.3.23}$$

for d_{xz}-d_{yz} orbitals. In both cases, the eigenvectors can be defined as

$$\begin{aligned} |\Gamma_1\rangle &= (|\uparrow\downarrow, 0\rangle + |0, \uparrow\downarrow\rangle) 2^{-1/2} & U_{\Gamma_1} &= U + J \\ |\Gamma_2\rangle &= |\uparrow, \uparrow\rangle & U_{\Gamma_2} &= U - 3J \\ |\Gamma_3\rangle &= (|\uparrow, \downarrow\rangle + |\downarrow, \uparrow\rangle) 2^{-1/2} & U_{\Gamma_3} &= U - 3J \\ |\Gamma_4\rangle &= (|\uparrow, \downarrow\rangle - |\downarrow, \uparrow\rangle) 2^{-1/2} & U_{\Gamma_4} &= U - J \\ |\Gamma_5\rangle &= |\downarrow, \downarrow\rangle & U_{\Gamma_5} &= U - 3J \\ |\Gamma_6\rangle &= (|\uparrow\downarrow, 0\rangle - |0, \uparrow\downarrow\rangle) 2^{-1/2} & U_{\Gamma_6} &= U - J \; . \end{aligned} \tag{A.3.24}$$

A.3.3 Broken local symmetries

In this section we show that the expectation values of $|\Gamma_4\rangle\langle\Gamma_4|$ and $|\Gamma_6\rangle\langle\Gamma_6|$ are different in the Gutzwiller wave function, despite the fact that their atomic energies are the same, see Eq. (A.3.24). This effect is in accordance with the symmetries of the exact ground state obtained from exact diagonalization of a finite system. The expansion to trivial order already shows a broken symmetry after an optimization of the variational parameters. However, we use the first-order approximation to illustrate why the expectation values are not the same.

Let us assume that both expectation values are equal. Then, both expectation values must lead to the same polynomials in $\rho_{ij}^{\sigma\sigma'}$ for arbitrary entries in $\rho_{ij}^{\sigma\sigma'}$ (as long as the p_x-p_y symmetry is obeyed). The key point of this argumentation is that the first-order diagram with four lines leads to different polynomials in $\rho_{ij}^{\sigma\sigma'}$. These differences cannot be compensated by any other contribution so that we can restrict our analysis to the external nodes with four lines. Furthermore, we can assume that the density matrix does not allow hopping processes between the orbitals. Then, the Γ_4 projector leads to external nodes that consist of the Graßmann variables

$$(a) = \tilde{c}_{u_1}^{\dagger}\tilde{c}_{d_1}^{\dagger}\tilde{c}_{d_1}\tilde{c}_{u_1} \quad \text{and} \quad (b) = \tilde{c}_{u_1}^{\dagger}\tilde{c}_{d_1}^{\dagger}\tilde{c}_{d_2}\tilde{c}_{u_2} \; . \tag{A.3.25}$$

The Γ_6 projector will lead to external nodes that consist of the Graßmann variables

$$(c) = \tilde{c}_{u_1}^{\dagger}\tilde{c}_{d_2}^{\dagger}\tilde{c}_{d_2}\tilde{c}_{u_1} \quad \text{and} \quad (d) = \tilde{c}_{u_1}^{\dagger}\tilde{c}_{d_2}^{\dagger}\tilde{c}_{u_2}\tilde{c}_{d_1} \; . \tag{A.3.26}$$

In first order, the internal nodes must be identical to the external nodes because we can contract only operators with the same spin-band index. The terms (b) and (d) lead to contributions that are proportional to $(\rho_x^{11})^2(\rho_x^{22})^2$. The diagrammatic expansion of Γ_4 includes the term (a) that leads to contributions that are proportional to $(\rho_x^{11})^4$. In contrast to that, the expansion of Γ_6 includes the term (b) which will lead to contributions that are proportional to $(\rho_x^{11})^2(\rho_x^{22})^2$. This explains the differences in the expectation values for the two projectors. The same argumentation can be used for higher orders.

A.3.4 Particle hole symmetry

The Hamilton operator displays particle-hole symmetry. This can be seen by mapping the particle operators to hole operators

$$\hat{c}_{i,\sigma}^{\dagger} \rightarrow (-1)^{||i||_1}\hat{h}_{i,\sigma} \; , \quad \hat{c}_{i,\sigma} \rightarrow (-1)^{||i||_1}\hat{h}_{i,\sigma}^{\dagger} \; , \tag{A.3.27}$$

where $||.||_1$ gives the one-norm. After we apply the transformation in Eq. (A.3.27) the sign of the hopping amplitudes has to be adapted

$$\hat{T}_{\mathrm{h}} = \sum_{i,j,\sigma,\sigma'} (-1)^{||i-j||_1} t_{ij}^{\sigma\sigma'} \hat{h}_{i,\sigma}^{\dagger}\hat{h}_{j,\sigma'} \; . \tag{A.3.28}$$

The transformed potential energy becomes

$$\hat{U}_{\mathrm{h}} = \sum_{l,\sigma_1,\ldots,\sigma_4} U_{\sigma_1\sigma_2\sigma_3\sigma_4}\hat{h}_{i,\sigma_1}^{\dagger}\hat{h}_{i,\sigma_2}^{\dagger}\hat{h}_{i,\sigma_3}\hat{h}_{i,\sigma_4} - (3U-5J)\sum_{i,\sigma}\hat{n}_{\sigma} + 2(3U-5J) \tag{A.3.29}$$

with $U_{\sigma_1\sigma_2\sigma_3\sigma_4}$ given as in section 3.1. We define the new vacuum state as the fully occupied state

$$|\mathrm{vac}\rangle_\mathrm{h} = \prod_{k,\gamma} \hat{c}^\dagger_{k,\gamma} |\mathrm{vac}\rangle \,. \tag{A.3.30}$$

Then, the new ground state of $\hat{T}_\mathrm{h}$ with density $n_0^\mathrm{h} = 1 - n_0$ can be written as

$$|\Psi_0\rangle_\mathrm{h} = \prod_{k,\gamma}^{\text{unocc.}} \hat{h}^\dagger_{k+Q,\gamma} |\mathrm{vac}\rangle_\mathrm{h} \,, \tag{A.3.31}$$

where the product runs over all unoccupied states of the ground state $|\Psi_0\rangle$ of $\hat{T}$ and $Q = (\pi, \pi)$. The numerical optimization of the Gutzwiller wave function can become very hard for densities above half filling. In these cases, we can optimize the Gutzwiller wave function of the transformed Hamilton operator with density $n_0^\mathrm{h} = 1 - n_0$ instead. The Fermi surface can be reconstructed from the mapping described above.

Nevertheless, the particle-hole mapping does not have to lead to the same numerical results because the chemical potential term is not explicitly taken into account in our calculations. This means, that we do not carry out the diagrammatic expansion for these terms. Instead, we drop the expectation values of $(3U - 5J) \sum_{i,\sigma} \langle \Psi_G | \hat{n}_\sigma | \Psi_g \rangle$ and fix the density by an appropriate choice of the Fermi energy. However, this changes the gradient $\partial_{\rho_{ij}^{\sigma\sigma'}} E_G$ of the Gutzwiller energy with respect to the density matrix. This mildly affects the optimization algorithm defined in subsection 3.3.1. For example, in some cases the deformed Fermi surfaces of both systems at half filling will not overlap perfectly.

A.4 Fermi surface optimization

A.4.1 Substitution of the external nodes with two lines

The particle-number conservation allows us to replace all diagrams that are connected to a single external node with two lines. Instead, we can express these diagrams by those which have four or more lines. In our expansion, the expectation value of the total particle number operator is given by

$$N_0 = \langle \hat{N} \rangle_G = M_\varnothing(\hat{N}) + \sum_{I_1,I_2}{}' \left(M^2_{I_1,I_2}(\hat{N}) + \sum_{m \geqslant 4} M^m_{I_1,I_2}(\hat{N}) \right) R_{I_1,I_2}, \tag{A.4.1}$$

where the coefficients $M^m_{I_1,I_2}(\hat{N})$ give the external weights with $|I_1| = m = |I_2|$. The value of the internal nodes and the line weight of all possible diagrams is given by R_{I_1,I_2}. Furthermore, the gauge constraints in Eq. (2.4.1) will guarantee that $M_\varnothing(\hat{N} = N_0) = N_0$. In the spin-band symmetric case for example, the coefficients $M^m_{I_1,I_2}(\hat{N})$ are diagonal and all equal. Then we can solve Eq. (A.4.1) for $R_{u1,u1} =$

$R_{d1,d1} = R_{u2,u2} = R_{d2,d2}$. Now, we can express all two-line diagrams in the potential energy with the value of $R_{u1,u1}$. This trick has already been used in [41] and [13]. For example, it can be used to ensure particle number conservation or to improve the convergence with respect to the cutoff length of the density matrix. When the external and internal node of a diagram are connected by a single line we have to consider the lattice alignments in which these nodes are far away. However, we may neglect the long-range diagrams if we use only external nodes with a larger number of lines. Nevertheless, it should be considered that the number of diagrams increases with the number of possible line combinations. For this reason, the number of all diagrams which are attached to an external node with 8 lines is much higher than the number of diagrams attached to the external node with two lines.

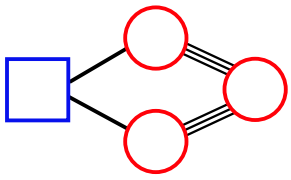
The first diagrams that have an external node with two lines occur in the second order of our expansion. At half filling it is possible to show that these diagrams have to vanish identically for our setup. This result is a direct implication of the idempotency of the density matrix which is shown as follows. We merge the spin-band and site index and consider the sum over all lattice sites. The external nodes $M^2_{I_1,I_2}$ are diagonal in the spin orbit index, so that we can write the sum of these diagrams as

$$\mathrm{ext}_2 = \sum_i M^2_{ii} \sum_{k,l} \tilde{\rho}_{ik} \tilde{\rho}_{il} R'_{kl} \ , \tag{A.4.2}$$

with $\tilde{\rho}_{ik}\tilde{\rho}_{il}R'_{kl} = R_{ii}$. The external coefficients M^2_{ii} do not depend on the index i because of the translational invariance and spin-band symmetry. When we express $\tilde{\rho}$ by the original density matrix ρ and use $\sum_i \rho_{ik}\rho_{il} = \rho_{kl}$, we obtain

$$\mathrm{ext}_2 = M \sum_{k,l} R_{kl}(\rho_{kl}(1-2n_0) + \delta_{lk} n_0^2) \ . \tag{A.4.3}$$

The external node is connected to two internal nodes that are connected to each other by 4, 6, or 8 lines. Therefore, ext_2 must vanish because $n_0 = 0.5$ and $R_{kk} = 0$. In third order, however, R_{kk} can take finite values as illustrated in the following diagram



.

The summation over i will let the two internal nodes collapse on the external node again. Then, the third internal node is connected to the external node by 6 lines and has a finite weight.

In the single-band case, the weight of the external nodes with four lines is zero. Therefore, setting the external nodes with two lines to zero gives the exact result of the diagrammatic expansion in all orders. In the multi-band case, however, this is not the case and the densities will not be conserved in a finite-order approximation.

In fact, the expectation value of the Gutzwiller wave function in Fig. A.6 at half filling give

$$\langle \hat{n}_{u1} \rangle_G = 0.49984 \,, \tag{A.4.4}$$

$$\langle \hat{n}_{u1} \rangle_G |_{\mathrm{ext}_2=0} = 0.49969 \,. \tag{A.4.5}$$

Note, that the idempotency cannot be incorporated properly into the optimization algorithm of the SPPS $|\Psi_0\rangle$. The ext_2 diagrams at half filling will lead to a finite contribution to the gradient $\partial_{\rho_{ij}} \langle \hat{H} \rangle$. This will be the case even when all long-range contributions are taken into account so that their weight vanishes but not their gradient. Thus, the optimization algorithm is affected by these diagrams and their exclusion may change the result of the optimization.

A.4.2 Cutoff length and long-range diagrams

In this section, we give a detailed discussion of the effect of the cutoff parameter dmax on the diagrammatic analysis. Furthermore, we discuss an improvement of the expansion by an elimination of some long-ranged diagrams, as discussed in section A.4.1.

First, we analyze the Fermi-surface deformations presented in section 3.3.2. A comparison with the result for dmax = 4 is shown in on the left side of Fig. A.6. Slight deviations are still visible for the inner Fermi surface. Both results of the outer Fermi surface lie on top of each other. At half filling, some long-ranged diagrams can be eliminated by setting the weight of the external nodes with two lines to zero. Both results agree very well as shown in the right panel of Fig. A.6.

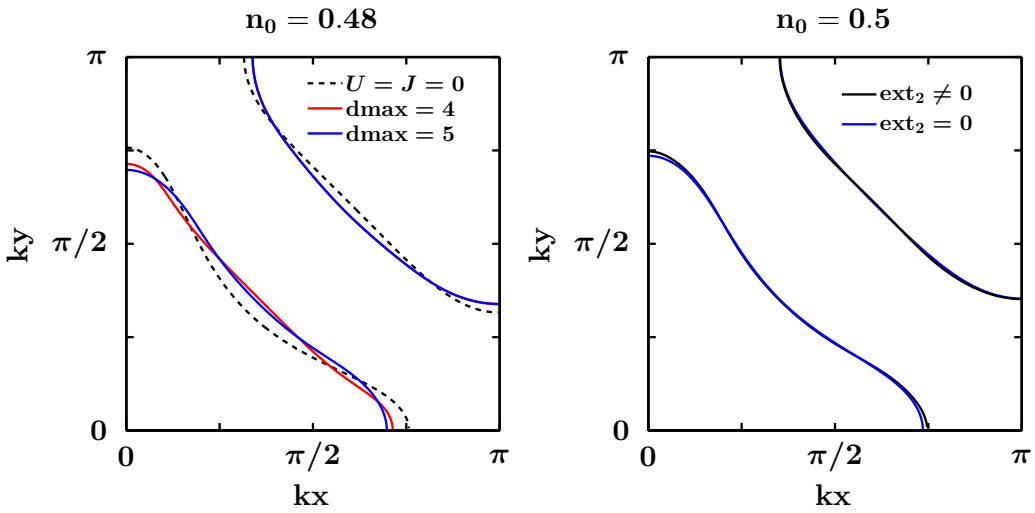


Figure A.6: Fermi surfaces for $U = 6.0$ and $J = 0.8$. Left: Fermi edges for dmax = 4 (red) and dmax = 5 (blue). The dashed black line gives the initial Fermi surface. Right: The Fermi surfaces at half filling. The blue line is obtained in the calculation scheme in which all the external nodes with two lines are explicitly set to zero. Both results agree very well. Both results are obtained for dmax = 5.

In the case of strong Fermi-surface deformations, as presented in section 3.3.3, the diagrammatic expansion depends more sensitively on the cutoff parameter dmax. A reduction of the cut-off length to dmax = 4 leads to a Fermi edge which is strongly curved as shown in the left panel of Fig. A.7. Again, the outer Fermi edge is more robust. At half filling, some long-ranged diagrams can be eliminated by setting the weight of the external nodes with two lines to zero as shown in In Fig. A.6. In this case, the Fermi edge is more straight but the overall deviations are only small compared to the differences to the initial Fermi surface.

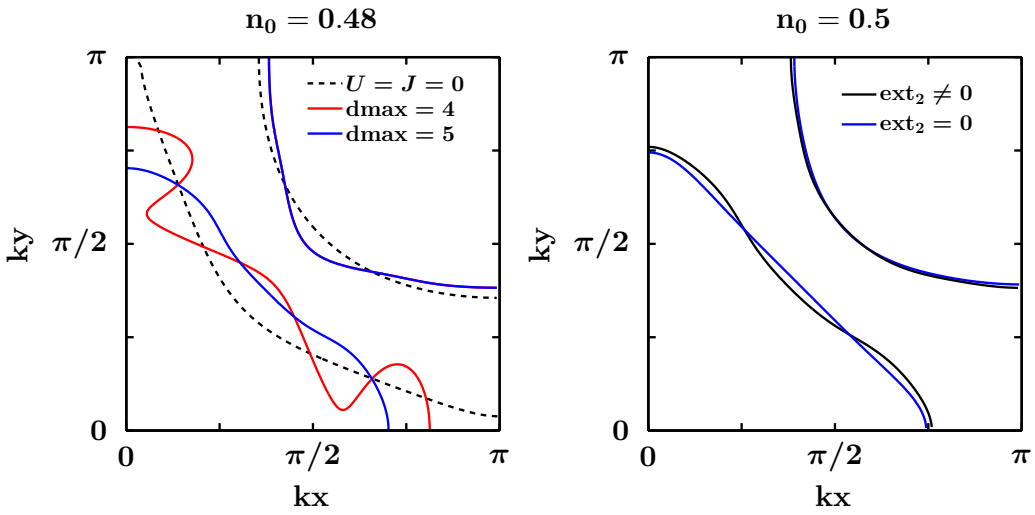


Figure A.7: Fermi surfaces at half filling and interaction strengths $U = 6.0$ and $J = 0.8$. Left: Fermi edges for dmax = 4 (red) and dmax = 5 (blue). The dashed black line gives the initial Fermi surface. Right: The blue line is obtained in the calculation scheme in which the external nodes of the potential energy are expressed through higher-order nodes. The blue edge for $n_k = 2$ seems less curved.

The left panel of Fig. A.8 shows the Fermi surface deformations for $n_0 = 0.48$ with different cutoff lengths, dmax = 4 and dmax = 5. The deviations between the Fermi edges are still large but the Pomeranchuk instability is present in both approximations. The right panel of Fig. A.8 shows the results with different cut-off parameters of the density matrix. The results for dmax = 5 and $r_{ij}^2 \leqslant 10$ lie almost on top of each other while the results obtained with dmax = 4 still deviates visibly, which might be caused by oscillations in ρ_{ij}. Therefore, the quality of the results can still be improved by an extension of the cutoff length.

We conclude that the effect of the cut-off parameter on the Fermi surface is still visible when the deformations change the symmetry or topology of the Fermi surface. Nevertheless, these qualitative features are not affected by the deviations. Furthermore, the exclusion of the diagrams attached to the external node with two lines can improve the convergence with respect to the cut-off parameter dmax. Additional strategies to incorporate long-ranged diagrams are given in section A.2.

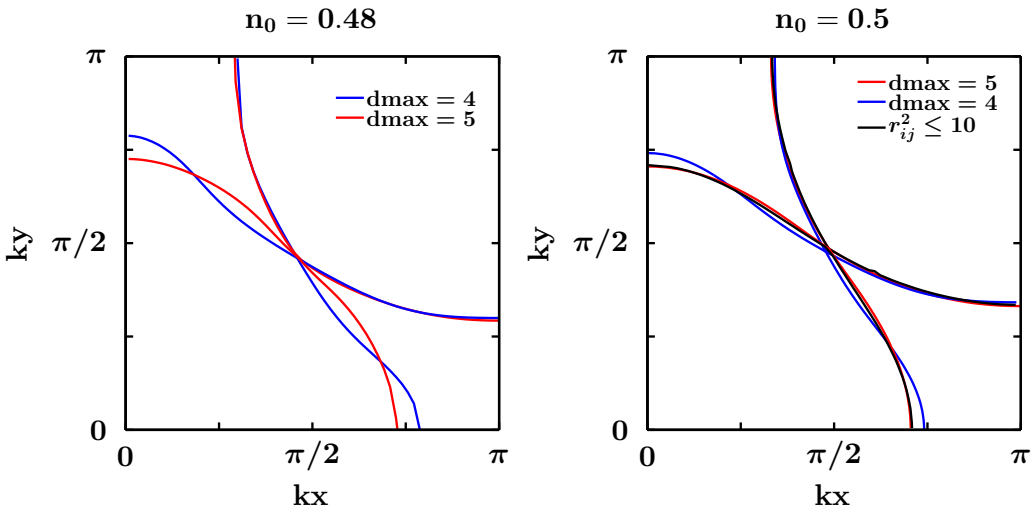


Figure A.8: Fermi surfaces for $U = 10$ and densities close to half band-filling. Left: The deviations between the cutoff length dmax = 4 and dmax=5 are considerable for $n_0 = 0.48$. Right: The deviations between the cutoff length dmax = 4 and dmax=5 are clearly discernible for $n_0 = 0.5$. The weights of the external nodes with two lines are set to zero.

Bibliography

[1] D. Pines and P. Nozières. *Theory of Quantum Liquids: Normal Fermi Liquids.* Advanced book classics. Addison-Wesley Publishing Company, Advanced Book Program, 1989.

[2] W. Nolting. *Grundkurs Theoretische Physik 7: Viel-Teilchen-Theorie.* Springer-Lehrbuch. Springer Berlin Heidelberg, 2014.

[3] M. C. Gutzwiller. *Effect of Correlation on the Ferromagnetism of Transition Metals.* Physical Review, 10:159, 1963.

[4] W. Metzner and D. Vollhardt. *Ground-state properties of correlated fermions: Exact analytic results for the Gutzwiller wave function.* Physical Review Letters, 59:121, 1987.

[5] W. Metzner and D. Vollhardt. *Analytic calculation of ground-state properties of correlated fermions with the Gutzwiller wave function.* Physical Review B, 37:7382, 1988.

[6] W. Metzner and D. Vollhardt. *Correlated lattice fermions in $d = \infty$ dimensions.* Physical Review Letters, 62:324, 1989.

[7] M. Kollar and D. Vollhardt. *Exact analytic results for the Gutzwiller wave function with finite magnetization.* Physical Review B, 65:155121, 2002.

[8] F. Gebhard. *Gutzwiller-korrelierte Wellenfunktionen in endlichen Dimensionen d: Eine systematische Entwicklung in $1/d$.* PhD thesis, RWTH Aachen, 1990.

[9] F. Gebhard. *Gutzwiller correlated wave functions in finite dimensions d: A systematic expansion in 1/d.* Physical Review B, 41:9452, 1990.

[10] J. Bünemann, F. Gebhard, and W. Weber. *Gutzwiller-correlated wave functions for degenerate bands: exact results in infinite dimensions.* J. Phys. Cond. Matt., 9:7343, 1997.

[11] J. Bünemann, F. Gebhard, T. Ohm, R. Umstaetter, S. Weiser, W. Weber, R. Claessen, D. Ehm, A. Harasawa, A. Kakizaki, A. Kimura, G. Nicolay, S. Shin, and V. N. Strocov. *Atomic correlations in itinerant ferromagnets: quasi-particle bands of nickel.* EPL (Europhysics Letters), 61:667 , 2003.

[12] T. Schickling J. Bünemann F. Gebhard and W. Weber. *Gutzwiller Density Functional Theory : a formal derivation and application to ferromagnetic nickel.* New Journal of Physics, 16:093034, 2014.

[13] J. Bünemann, T. Schickling, and F. Gebhard. *Variational study of Fermi surface deformations in Hubbard models.* EPL (Europhysics Letters), 98:27006, 2012.

[14] J. Bünemann. *The Gutzwiller Variational theory and related Methods for Correlated Electron Systems.* Habilitation thesis, Philipps-Universität Marburg, 2009.

[15] A. L. Fetter and J. D. Walecka. *Quantum Theory of Many-particle Systems.* Dover Books on Physics. Dover Publications, Mineola New-York, 2003.

[16] A. Rapp, W. Hofstetter, and G. Zaránd. *Trionic phase of ultracold fermions in an optical lattice: A variational study.* Physical Review B, 77:144520, 2008.

[17] J. Stoer and R. Bulirsch. *Introduction to Numerical Analysis.* Texts in Applied Mathematics. Springer Berlin Heidelberg, 2002.

[18] T. Müller, S. Fölling, A. Widera, and I. Bloch. *State preparation and dynamics of ultracold atoms in higher lattice orbitals.* Physical Review Letters, 99:200405, 2007.

[19] M. Ölschläger, T. Kock, and G. Wirth. *Interaction-induced chiral $p_x \pm ip_y$ superfluid order of bosons in an optical lattice.* New Journal of Physics, 15:83041, 2013.

[20] P. Soltan-Panahi, D.-S. Lühmann, J. Struck, P. Windpassinger, and K. Sengstock. *Quantum phase transition to unconventional multi-orbital superfluidity in optical lattices.* Nature Physics, 8:71, 2012.

[21] Y. Li, E. H. Lieb, and C. Wu. *Exact results for itinerant ferromagnetism in multiorbital systems on square and cubic lattices.* Physical Review Letters, 112:217201, 2014.

[22] D. Baeriswyl, D. Eichenberger, and M. Menteshashvili. *Variational ground states of the two-dimensional Hubbard model.* New Journal of Physics, 11:075010, 2009.

[23] L. F. Tocchio, F. Becca, A. Parola, and S. Sorella. *Role of backflow correlations for the nonmagnetic phase of the t- t′ Hubbard model.* Physical Review B, 78:041101, 2008.

[24] L. F. Tocchio, F. Becca, and C. Gros. *Backflow correlations in the Hubbard model: An efficient tool for the study of the metal-insulator transition and the large-U limit.* Physical Review B, 83:195138, 2011.

[25] L. F. Tocchio, F. Becca, and C. Gros. *Strong renormalization of the Fermi-surface topology close to the Mott transition.* Physical Review B, 86:035102, 2012.

[26] R. Rüger, L. F. Tocchio, R. Valentí, and C. Gros. *The phase diagram of the square lattice bilayer Hubbard model: A variational Monte Carlo study.* New Journal of Physics, 16:033010, 2014.

[27] S. Sugano, T. Yukito, and H. Kamimura. *Multiplets of transition metal ions in crystals.* Academic Press, New York, London, 1970.

[28] J. Bünemann, F. Gebhard, T. Schickling, and W. Weber. *Numerical minimisation of Gutzwiller energy functionals.* Physica Status Solidi B, 249:1282, 2012.

[29] T. Schickling, F. Gebhard, J. Bünemann, L. Boeri, O. K. Andersen, and W. Weber. *Gutzwiller theory of band magnetism in LaOFeAs.* Physical Review Letters, 108:036406, 2012.

[30] J. H. Van Vleck. *Models of exchange coupling in ferromagnetic media* Rev. Mod. Phys., 25:220, 1953.

[31] I. J. Pomeranchuk. JETP, 8:361, 1958.

[32] W. Metzner, D. Rohe, and S. Andergassen. *Soft fermi surfaces and breakdown of Fermi-liquid behavior.* Physical Review Letters, 91:066402, 2003.

[33] H. Yamase, V. Oganesyan, and W. Metzner. *Mean-field theory for symmetry-breaking Fermi surface deformations on a square lattice.* Physical Review B, 72:035114, 2005.

[34] W. Metzner, L. Dell'Anna, and H. Yamase. *Nematic order and non-Fermi liquid behavior from a Pomeranchuk instability in a two-dimensional electron system.* Journal of Physics: Conference Series, 150:032058, 2009.

[35] J. Kaczmarczyk, J. Spałek, T. Schickling, and J. Bünemann. *Superconductivity in the two-dimensional Hubbard model: Gutzwiller wave function solution.* Physical Review B, 88:115127, 2013.

[36] J. Kaczmarczyk. *Private communications*, 2015.

[37] M. Galassi et al. *GNU Scientific Library Reference Manual.* 3rd edition.

[38] G. B. Arfken, H. J. Weber, and F. E. Harris. *Mathematical Methods for Physicists: A Comprehensive Guide.* Elsevier, 2012.

[39] I. Wolfram Research. *Mathematica.* Wolfram Research, Inc., Champaign, Illinois, Version 9, 2012.

[40] R. Alsfasser and E. Riedel. *Moderne anorganische Chemie: mit CD-ROM.* de Gruyter, 2007.

[41] W. Metzner. *Variational theory for correlated lattice fermions in high dimensions.* Zeitschrift für Physik B Condensed Matter, 77:253, 1989.

Danksagung

Hiermit möchte ich mich bei Prof. Dr. Florian Gebhard für die Betreuung und sorgfältige Korrektur meiner Doktorarbeit bedanken. Des Weiteren bedanke ich mich bei Priv.-Doz. Dr. Jörg Bünemann, der die Grundsteine dieses Projekts gelegt hat. Prof. Dr. Peter Lenz möchte ich für die Übernahme des Zweitgutachtens danken.

Besonderen Dank gilt Dr. Jan Kaczmarczyk, der mir seine Ergebnisse zur diagrammatischen Analyse des Einband-Hubbard-Modells, siehe Abschnitt 3.3.4, zur Verfügung gestellt hat.

Weiterhin möchte ich mich bei allen Mitgliedern unserer Arbeitsgruppe bedanken, die mir mit guten Ratschlägen oder nützlichen Ideen weitergeholfen haben.

Ich bedanke mich bei meiner Famillie und bei meinen Freunden für die persönliche Unterstützung, die ich in den letzten Jahren erfahren habe.